

Approche temps-fréquence pour la séparation aveugle de sources non-stationnaires

Cédric Févotte

Thèse de Doctorat
de l'École Centrale de Nantes et de l'Université de Nantes

Soutenue le 23 octobre 2003 devant un jury composé de :

- | | |
|---|---------------------------|
| – Jean-François Cardoso , Directeur de Recherche CNRS, ENST Paris, | président du jury |
| – Christian Jutten , Professeur, INPG Grenoble, | rapporteur |
| – Yannick Deville , Professeur, Université Paul Sabatier Toulouse, | rapporteur |
| – Lieven De Lathauwer , Chargé de Recherche CNRS, ENSEA Cergy-Pontoise | |
| – Éric Le Carpentier , Maître de Conférences, Ecole Centrale de Nantes | |
| – Christian Doncarli , Professeur, Ecole Centrale de Nantes, | directeur de thèse |

Résumé :

Le travail présenté dans cette thèse concerne la Séparation Aveugle de Sources (SAS). La SAS consiste à estimer n signaux inconnus (les sources) d'après la seule connaissance de m mélanges de ces signaux (les observations).

Nous étudions dans un premier temps l'identifiabilité des sources dans le problème de la séparation aveugle de mélanges linéaires instantanés (sur-)déterminés ($m \geq n$). Plusieurs modèles de sources identifiables sont examinés. En particulier, il est montré que si les sources possèdent une diversité temporelle et/ou fréquentielle, leur indépendance mutuelle à l'ordre 2 suffit à garantir leur identifiabilité.

Nous décrivons dans un deuxième temps quelques méthodes de SAS construites sur les hypothèses des modèles identifiables examinés. Ces méthodes ont en commun d'être basées sur une procédure de diagonalisation simultanée de matrices. Nous nous intéressons plus particulièrement à une méthode adaptée aux sources non-stationnaires, basée sur la diagonalisation simultanée de matrices extraites du Spectre de Wigner-Ville Spatial (SWVS) des observations à certains points du plan temps-fréquence (et après blanchiment spatial). Nous proposons une approche théorique dans un contexte stochastique qui justifie l'approximation en pratique du SWVS par des Représentations Temps-Fréquence Spatiales de la classe de Cohen. Les performances de la méthode dépendent dans une large mesure de la sélection des points temps-fréquence et nous proposons un critère de sélection robuste basé sur la détection d'auto-termes simples des sources. Une étude statistique des performances des méthodes exposées est réalisée sur des signaux synthétiques non-stationnaires de type TVARMA. Une série de critères d'évaluation originaux est proposée.

Enfin, dans un troisième temps, nous montrons comment les méthodes de SAS présentées dans le cas de mélanges linéaires instantanés peuvent être généralisées à la séparation de mélanges convolutifs à Réponse Impulsionnelle Finie en utilisant une procédure de *bloc*-diagonalisation simultanée. Quelques résultats sur des mélanges synthétiques de signaux audio sont présentés.

Mots clés :

Séparation Aveugle de Sources, Spectre de Wigner-Ville Spatial, Représentations Temps-Fréquence Spatiales, Signaux Non-Stationnaires.

À mes parents

« L'homme de Science le sait bien, lui, que, sans la Science, l'homme ne serait qu'un stupide animal sottement occupé à s'adonner aux vains plaisirs de l'amour dans les folles prairies de l'insouciance, alors que la Science, et la Science seule, a pu, patiemment, au fil des siècles, lui apporter l'horloge pointeuse et le parcmètre automatique sans lesquels il n'est pas de bonheur terrestre possible. »

Pierre Desproges, *Vivons heureux en attendant la mort*

Table des matières

Remerciements	11
Introduction	13
1 Identifiabilité des mélanges linéaires instantanés	19
1.1 Généralités	19
1.1.1 Hypothèses fondamentales	19
1.1.2 Indéterminations	20
1.1.3 Identifiabilité	21
1.2 Orthonormalisation	22
1.3 Modèles identifiables	24
1.3.1 Sources non-stationnaires colorées et mutuellement décorréliées (modèle \mathbb{M}_1)	25
1.3.2 Sources non-stationnaires blanches et mutuellement décorréliées (modèle \mathbb{M}_2)	27
1.3.3 Sources stationnaires colorées et mutuellement décorréliées (modèle \mathbb{M}_3) .	27
1.3.4 Sources i.i.d et mutuellement indépendantes (modèle \mathbb{M}_4)	28
1.3.5 Discussion	29
2 Algorithmes standard de séparation en mélange linéaire instantané	31
2.1 Séparation de sources i.i.d (modèle \mathbb{M}_4)	31
2.1.1 Indépendance et décorrélation	32
2.1.2 Quelques méthodes d'ACI en théorie	34
2.1.3 Quelques méthodes d'ACI en pratique	39
2.2 Séparation de sources stationnaires colorées (modèle \mathbb{M}_3)	42
2.2.1 SOBI en théorie	43
2.2.2 SOBI en pratique	44
2.2.3 Discussion	45
3 Séparation de sources non-stationnaires en mélange linéaire instantané	47
3.1 Séparation de sources non-stationnaires blanches (modèle \mathbb{M}_2)	47
3.2 Contributions à une approche temps-fréquence (modèle \mathbb{M}_1)	51
3.2.1 Méthode théorique	52
3.2.2 Implémentation pratique	53
3.2.3 Sélection des points temps-fréquence	54
3.3 Résultats comparatifs sur des mélanges synthétiques	58
3.3.1 Critères d'évaluation	58
3.3.2 Performances comparées	64
3.4 Conclusion	79

4	Généralisation aux mélanges convolutifs RIF	81
4.1	Modèle et notations	82
4.1.1	Retour à un mélange linéaire instantané	82
4.1.2	Discussion	83
4.2	Méthode générale	84
4.2.1	Représentation temps-retard	84
4.2.2	Orthonormalisation	85
4.2.3	Rotation	85
4.2.4	Estimation des sources	86
4.2.5	Discussion	86
4.3	SOBI convolutif	87
4.4	Quelle implémentation pour des sources non-stationnaires ?	89
4.4.1	L'approche temps-fréquence	89
4.4.2	Sélection des points t-f	90
4.5	Conclusion	92
	Conclusion générale	95
A	Diagonalisation et bloc-diagonalisation simultanée de matrices	97
A.1	Diagonalisation approchée	97
A.1.1	Base orthonormale complexe	97
A.1.2	Base orthonormale réelle	102
A.2	Diagonalisation simultanée approchée	103
A.2.1	Base orthonormale complexe	103
A.2.2	Base orthonormale réelle	105
A.3	Bloc-diagonalisation approchée	105
A.3.1	Base orthonormale complexe	105
A.3.2	Matrices réelles	107
A.4	Bloc-diagonalisation simultanée approchée	109
A.4.1	Base orthonormale complexe	109
A.4.2	Matrices réelles	110
B	Déconvolution aveugle	113
B.1	Modèle	113
B.2	Notations	113
B.3	Méthode	114
B.4	Identifiabilité	115

Remerciements

Mes premiers remerciements vont à l'équipe ADTS de l'IRCCyN pour m'avoir permis de travailler à ma thèse dans d'*excellentes* conditions. Christian, merci infiniment pour ton soutien scientifique et moral ; plus encore, merci de m'avoir fait découvrir Sting version jazz et Jimmy Smith avec Wes Montgomery. Marie, ces remerciements prendraient plus de place que le contenu scientifique de ce mémoire s'il me fallait écrire ici tout le bien que je pense de toi. Alexandra, un grand merci pour ton aide.

Merci au GdR-ISIS de m'avoir permis d'effectuer un séjour de recherche en Slovénie et un déplacement au Japon. Clin d'oeil à mes complices de l'Action Jeunes Chercheurs «Ressources pour la séparation de sources audio», Rémi Gribonval et Emmanuel Vincent ; les nombreuses discussions scientifiques (souvent animées !) que nous avons eues m'ont beaucoup apporté.

Enfin, merci à mes éminents membres du jury, notamment aux rapporteurs, de l'intérêt qu'ils ont bien voulu porter à mon travail.

Chers acolytes des pauses café et compagnons de labeur, je vous tire ma révérence. Merci d'avoir tant contribué à cette agréable atmosphère de travail. Mentions spéciales à Pierre-Long Beach-Emmanuel, *Manhattan* Fél, Fréd le loup, to be Franck, mon ami Pierrot, au Claire de la lune, les bains de minuit à Noirmoutier, les bains de 7h à St-Molf.

Prix des filles jeunes à Benjamin pour l'ensemble de son œuvre.

Lucas, vieux compère, j'ai une pensée pour ce mémorable jour des résultats du bac. Bruno-Edouard, les maths de prépa me servent enfin à quelque chose...

Merci aux développeurs des logiciels libres que j'ai utilisés pour mes travaux : XEmacs, L^AT_EX, Prosper, Mozilla, OpenOffice et surtout XBlast.

Enfin, pour leur soutien ou réconfort, je souhaite remercier : Pauline et la piscine Léo Lagrange, les boulangères de la rue de la Fosse, les Little Rabbits, mon public, mes fans, et la merveilleuse ville de Nantes, pour l'ensemble de son œuvre, également.

So long !

Mont-sur-Meurthe, 25 décembre 2003

Introduction

La Séparation Aveugle de Sources (SAS) consiste à estimer n signaux inconnus (les sources) d'après la seule connaissance de m mélanges de ces signaux (les observations). Le terme *aveugle* signifie que les sources ne sont pas observées.

La SAS possède de nombreuses applications (voir [Dev01] pour un panorama), notamment dans les domaines du Traitement du Signal Audio et du Traitement du Signal Biomédical. Dans le domaine de l'Audio, la SAS concerne par exemple le «démixage» d'extraits musicaux stéréophoniques ou le rehaussement de la parole, pour la téléphonie mobile ou pour l'élaboration de prothèses auditives (voir [VFea03] pour un panorama des applications de la SAS en Audio). Dans le domaine biomédical, de nombreux travaux existent sur la séparation d'EléctroEncéphaloGrammes (EEG) [VO00], EléctroCardioGrammes (ECG) [LMV00] ou plus récemment Eléctro-MyoGrammes (EMG) [FFDM04, FLF⁺03] (voir [Dev03] pour un panorama des applications biomédicales de la SAS).

La SAS possède plusieurs degrés de difficulté, selon les caractéristiques des sources et surtout du mélange. Plusieurs paramètres sont à prendre en considération.

Le premier paramètre est le nombre m d'observations par rapport au nombre n de sources. Intuitivement, on conçoit aisément que le cas *(sur-)déterminé* ($m \geq n$) est plus simple à résoudre que le cas *sous-déterminé* ($m < n$), ce dernier cas ne pouvant généralement être résolu qu'au prix d'une importante information *a priori* sur les sources.

Un deuxième paramètre est la nature du mélange. Le mélange le plus simple est le mélange linéaire instantané : à chaque instant les observations sont des combinaisons linéaires des sources au même instant. Dans la réalité, les mélanges sont souvent convolutifs : à chaque instant les observations dépendent aussi des sources aux instants précédents. En Audio par exemple, lorsque deux musiciens (jouant d'un instrument acoustique) sont enregistrés par différents micros dans une salle, l'effet de réverbération de la salle produit au niveau des micros un mélange convolutif. Dans la réalité les mélanges sont aussi souvent variant dans le temps. Par exemple lorsque des musiciens se déplacent, les fonctions de transfert de la salle entre les émetteurs (les musiciens) et les récepteurs (les micros) changent suivant les positions des émetteurs et récepteurs, produisant donc sur les micros un mélange variant dans le temps. Enfin, les mélanges peuvent être non-linéaires, par exemple certains micros ont un comportement non-linéaire, bien souvent négligeable, par rapport à l'intensité des sources enregistrées.

Enfin, un troisième paramètre est la nature des sources. Nous verrons par la suite que la plupart des méthodes de SAS reposent sur l'hypothèse que les sources sont mutuellement indépendantes, au moins à l'ordre 2. Ceci est une contrainte forte qui ne peut pas toujours être vérifiée en pratique. Nous verrons aussi que le cas de sources non-stationnaires est plus difficile à «appréhender» que le cas de sources stationnaires.

La SAS est une discipline encore jeune, les premiers travaux sont attribués à Héroult & Jutten au milieu des années 80 [HJA85] (voir [JT00] pour une histoire de la SAS). De ce fait, seuls les mélanges Linéaires Invariant dans le Temps (LTI) ont été pour le moment essentiellement abordés, et plus particulièrement les mélanges linéaires instantanés. Quelques auteurs se sont

intéressés à la séparation de mélanges (post) non-linéaires [TJ99, EK02] ou variant dans le temps [Yer03].

Dans cette thèse nous nous sommes principalement intéressés à la séparation de mélanges linéaires instantanés, notamment de sources non-stationnaires. Comme dans la majorité des travaux sur la SAS, l'approche que nous adoptons est *stochastique*, c-à-d que les sources sont traitées comme les réalisations de *séquences aléatoires*. Les hypothèses que nous serons amenés à formuler sur les sources portent ainsi sur les séquences aléatoires qui les modélisent. Il est à noter que des approches «déterministes» de la SAS existent ; par exemple la séparation aveugle de sinusoides noyées dans du bruit fait l'objet de nombreux travaux d'Analyse Spectrale.

Nous étudions dans le chapitre 1 l'identifiabilité des mélanges linéaires instantanés d'une manière générale. Nous discutons des indéterminations de la SAS et de la notion d'identifiabilité. Nous verrons qu'il n'est possible de résoudre un problème de SAS qu'au prix de certaines contraintes sur les sources. Nous proposons alors quelques modèles de sources *identifiables*, c-à-d, répondant à des hypothèses suffisamment contraignantes pour rendre le problème soluble. Deux modèles généraux se distinguent. Si les sources possèdent une *diversité temporelle et/ou fréquentielle* alors leur indépendance mutuelle à l'ordre 2 suffit à leur identifiabilité. En revanche, si les sources ne possèdent pas de *diversité temporelle et/ou fréquentielle*, c-à-d, si elles sont modélisées par des séquences identiquement et indépendamment distribuées (i.i.d), dont au plus une gaussienne, nous verrons que leur indépendance mutuelle *stricte* est nécessaire.

Dans les chapitres 2 et 3, nous décrivons quelques méthodes de séparation construites sur les hypothèses des modèles de sources étudiés. Le chapitre 2 décrit quelques méthodes standard traitant le cas de sources stationnaires. Nous verrons notamment que dans le cas de sources i.i.d, le problème de la SAS peut être résolu par des méthodes d'Analyse en Composantes Indépendantes (ACI). Une attention particulière sera portée aux algorithmes JADE et SOBI, basés sur un processus de *diagonalisation simultanée* de matrices.

Le chapitre 3 est quant à lui dédié à la séparation de sources non-stationnaires. Une solution au sens du Maximum de Vraisemblance, proposée par Pham & Cardoso est étudiée dans la partie 3.1. Puis, nous étudions dans la partie 3.2 une méthode basée sur la diagonalisation simultanée de matrices extraites du Spectre de Wigner-Ville Spatial (SWVS) des observations à certains points du plan temps-fréquence (et après blanchiment spatial). Nous proposons une approche théorique dans un contexte stochastique qui justifie l'approximation en pratique du SWVS par des Représentations Temps-Fréquence Spatiales de la classe de Cohen. Les performances de la méthode dépendent dans une large mesure de la sélection des points temps-fréquence et nous proposons un critère de sélection robuste basé sur la détection d'auto-termes simples des sources. Une étude statistique des performances des méthodes exposées, réalisée sur des signaux synthétiques non-stationnaires de type TVARMA, est proposée dans la partie 3.3, où nous proposons en outre une série de critères d'évaluation originaux.

Nous montrons dans le chapitre 4 comment les méthodes de SAS présentées dans le cas de mélanges linéaires instantanés peuvent être généralisées à la séparation de mélanges convolutifs à Réponse Impulsionnelle Finie (RIF) en utilisant une procédure de *bloc-diagonalisation* simultanée. Nous proposons en annexe un algorithme de type Jacobi pour la bloc-diagonalisation simultanée. Cependant, cet algorithme soulève encore en pratique des problèmes de convergence dont nous discuterons. Quelques résultats sur des mélanges synthétiques de signaux audio sont présentés.

Enfin, plusieurs perspectives de recherche seront proposées en conclusion.

Dans un tout premier temps, nous définissons les notations et quantités mathématiques utilisées par la suite.

Notations matricielles

Soit une matrice $\mathbf{M} = \{m_{ij}\} \in \mathbb{C}^{m \times n}$. On note $\overline{\mathbf{M}}$ sa conjuguée, \mathbf{M}^T sa transposée et $\mathbf{M}^H = \overline{\mathbf{M}}^T$ sa transconjugée. \mathbf{M} est une matrice hermitienne si elle vérifie $\mathbf{M}^H = \mathbf{M}$.

La notation $\|\mathbf{M}\|_F$ désignera la norme de Frobenius définie par :

$$\|\mathbf{M}\|_F = \sqrt{\sum_{ij} |m_{ij}|^2} = \sqrt{\text{trace}(\mathbf{M}^H \mathbf{M})} \quad (1)$$

La norme p d'un vecteur \mathbf{m} de $\mathbb{C}^{n \times 1}$ ou $\mathbb{C}^{1 \times n}$, de composantes $\{m_1, \dots, m_n\}$, est définie par :

$$\|\mathbf{m}\|_p = \left(\sum_{i=1}^n |m_i|^p \right)^{\frac{1}{p}} \quad (2)$$

La notation $\text{diag}(\mathbf{m})$ désignera la matrice diagonale de $\mathbb{C}^{n \times n}$ dont les éléments diagonaux sont égaux aux composantes de \mathbf{m} . La notation $\text{diag}(\mathbf{M})$ désignera la matrice diagonale de $\mathbb{C}^{n \times n}$ dont les termes diagonaux sont égaux aux termes diagonaux de \mathbf{M} .

La matrice identité de dimension $m \times m$ sera notée \mathbf{I}_m . La matrice nulle de dimension $m \times m$ sera notée $\mathbf{0}_m$.

Pseudo-inverse de Moore-Penrose

La pseudo-inverse de Moore-Penrose de $\mathbf{M} \in \mathbb{C}^{m \times n}$ est l'unique matrice $\mathbf{M}^\# \in \mathbb{C}^{n \times m}$ vérifiant [GL96] :

$$\mathbf{M} \mathbf{M}^\# \mathbf{M} = \mathbf{M} \quad (3)$$

$$\mathbf{M}^\# \mathbf{M} \mathbf{M}^\# = \mathbf{M}^\# \quad (4)$$

$$(\mathbf{M} \mathbf{M}^\#)^H = \mathbf{M} \mathbf{M}^\# \quad (5)$$

$$(\mathbf{M}^\# \mathbf{M})^H = \mathbf{M}^\# \mathbf{M} \quad (6)$$

Si $m \geq n$ et si $(\mathbf{M}^H \mathbf{M})$ est inversible, alors :

$$\mathbf{M}^\# = (\mathbf{M}^H \mathbf{M})^{-1} \mathbf{M}^H \quad (7)$$

Si $m \leq n$ et si $(\mathbf{M} \mathbf{M}^H)$ est inversible, alors :

$$\mathbf{M}^\# = \mathbf{M}^H (\mathbf{M} \mathbf{M}^H)^{-1} \quad (8)$$

Si $m \geq n$ et si \mathbf{M} est de rang plein (égal à n), si $\mathbf{M} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}$ désigne une décomposition en valeurs singulières de \mathbf{M} telle que :

- $\mathbf{U} \in \mathbb{C}^{m \times n}$ vérifie $\mathbf{U}^H \mathbf{U} = \mathbf{I}_n$,
- $\mathbf{\Sigma} \in \mathbb{R}^{n \times n}$ est diagonale > 0 ,
- $\mathbf{V} \in \mathbb{C}^{n \times n}$ est orthonormale,

alors la pseudo-inverse de Moore-Penrose de \mathbf{M} s'écrit :

$$\mathbf{M}^\# = \mathbf{V}^H \mathbf{\Sigma}^{-1} \mathbf{U}^H \quad (9)$$

Matrices positives

Soit $\mathbf{M} \in \mathbb{C}^{n \times n}$. On dit que \mathbf{M} est *positive* (que l'on note $\mathbf{M} \geq 0$) ssi $\forall \mathbf{x} \in \mathbb{C}^{n \times 1} \setminus \{0\}$, $\mathbf{x}^H \mathbf{M} \mathbf{x} \geq 0$. On dit que \mathbf{M} est *définie positive* (que l'on note $\mathbf{M} > 0$) ssi $\forall \mathbf{x} \in \mathbb{C}^{n \times 1} \setminus \{0\}$, $\mathbf{x}^H \mathbf{M} \mathbf{x} > 0$.

Racine carrée d'une matrice hermitienne positive

Soit $\mathbf{M} \in \mathbb{C}^{n \times n}$ une matrice hermitienne positive. \mathbf{M} possède plusieurs *racines carrées*, c-à-d qu'il existe plusieurs matrices $\mathbf{X} \in \mathbb{C}^{n \times n}$ telles que $\mathbf{M} = \mathbf{X} \mathbf{X}^H$. Cependant, \mathbf{M} ne possède qu'une unique racine carrée *hermitienne positive* [GL96]. On notera $\mathbf{M}^{\frac{1}{2}}$ cette matrice, qui sera désignée comme *la* racine carrée de \mathbf{M} et telle que :

$$\mathbf{M} = \mathbf{M}^{\frac{1}{2}} \mathbf{M}^{\frac{H}{2}} = \mathbf{M}^{\frac{1}{2}} \mathbf{M}^{\frac{1}{2}} \quad (10)$$

Si \mathbf{M} est hermitienne définie positive, alors $(\mathbf{M}^{\frac{1}{2}})^{-1} = (\mathbf{M}^{-1})^{\frac{1}{2}}$, $(\mathbf{M}^{\frac{1}{2}})^H = (\mathbf{M}^H)^{\frac{1}{2}}$, $(\mathbf{M}^H)^{-1} = (\mathbf{M}^{-1})^H$, de sorte que la notation $\mathbf{M}^{-\frac{H}{2}}$, utilisée par la suite, ne présente aucune ambiguïté (les exposants «commutent»).

Séquences aléatoires

Soit $\{x[t]\}_{t=0 \dots T-1}$ une séquence aléatoire (éventuellement complexe). La notation $x[\cdot]$ désignera un échantillon quelconque de la séquence. Privé d'indice (temporel), x désigne l'ensemble des échantillons :

$$x \stackrel{\text{def}}{=} [x[0], \dots, x[T-1]]$$

La densité de probabilité de x est notée $p(x) = p(x[0], \dots, x[T-1])$. Sauf indication contraire, la densité de probabilité d'une variable aléatoire v sera notée $p(v)$, c-à-d que la variable de la fonction identifie la densité. Par exemple, si la séquence est indépendamment distribuée on écrira :

$$p(x) = \prod_{t=0}^{T-1} p(x[t]) \quad (11)$$

où $p(x[t])$ désigne alors la densité de probabilité de l'échantillon $x[t]$ à l'instant t .

Nous allons à présent définir la covariance et le Spectre de Wigner-Ville de séquences aléatoires. Afin de simplifier les définitions nous supposons que les séquences sont de longueur infinie. Une séquence x de longueur finie T pourra alors être représentée par une séquence de longueur infinie égale à x si $t \in \llbracket 0, T-1 \rrbracket$, nulle ailleurs.

Covariance

On note $r_{xx}[t, \tau]$ l'auto-covariance de la séquence x , supposée centrée, définie $\forall (t, \tau) \in \mathbb{Z}^2$ par :

$$r_{xx}[t, \tau] \stackrel{\text{def}}{=} \mathbb{E}\{x[t + \tau] \bar{x}[t]\} \quad (12)$$

où $\mathbb{E}\{\cdot\}$ désigne l'espérance d'une variable aléatoire. La covariance croisée de deux séquences aléatoires centrées x et y est notée :

$$r_{xy}[t, \tau] \stackrel{\text{def}}{=} \mathbb{E}\{x[t + \tau] \bar{y}[t]\} \quad (13)$$

Soit \mathbf{x} une séquence aléatoire vectorielle (c-à-d, $\forall t \in \mathbb{Z}, \mathbf{x}[t] \in \mathbb{C}^{m \times 1}$). En supposant \mathbf{x} centrée, la covariance de \mathbf{x} est définie par :

$$\mathcal{R}_{\mathbf{xx}}[t, \tau] \stackrel{\text{def}}{=} \mathbb{E}\{\mathbf{x}[t + \tau] \mathbf{x}[t]^H\} \quad (14)$$

Spectre de Wigner-Ville

On note $s_{xx}[t, f]$ le Spectre de Wigner-Ville (SWV) de la séquence x centrée, défini pour $(t, f) \in \mathbb{Z} \times [0, \frac{1}{2}[$ par [Fla99] :

$$s_{xx}[t, f] = 2 \sum_{\tau=-\infty}^{+\infty} \mathbb{E}\{x[t + \tau] \bar{x}[t - \tau]\} e^{-j4\pi f \tau} \quad (15)$$

$$= 2 \sum_{\tau} r_{xx}[t - \tau, 2\tau] e^{-j4\pi f \tau} \quad (16)$$

La notation $[t, f]$ signifie que le SWV est discret en temps et continu en fréquence. Le SWV est périodique en fréquence de période $\frac{1}{2}$. La notation $s_{xx}[t, f]$ désignera un échantillonnage à pas fixe du SWV de F points sur l'axe fréquentiel, pour $(t, f) \in \mathbb{Z} \times [0, F - 1]$:

$$s_{xx}[t, f] = s_{xx}\left[t, \frac{f}{2F}\right] \quad (17)$$

Si la séquence x est *stationnaire et colorée*, alors $r_{xx}[t, \tau] = r_{xx}[\tau]$ et :

$$s_{xx}[t, f] = 2 \sum_{\tau=-\infty}^{+\infty} r_{xx}[2\tau] e^{-j4\pi f \tau} \quad (18)$$

$$= 2 \sum_{\tau \text{ pair} \in \mathbb{Z}} r_{xx}[\tau] e^{-j2\pi f \tau} \quad (19)$$

$$= s_{xx}(f) \quad (20)$$

Pour des signaux stationnaires à temps continu le SWV est égal à la Densité Spectrale de Puissance (DSP) (voir [Fla99]). Ceci n'est pas vrai pour des signaux à temps discret. En effet, la DSP $dsp_{xx}(f)$ d'un signal stationnaire x à temp discret est définie comme la transformée de Fourier de sa fonction d'auto-covariance :

$$dsp_{xx}(f) = \sum_{\tau=-\infty}^{+\infty} r_{xx}[\tau] e^{-j2\pi f \tau} \quad (21)$$

On voit donc que par rapport à la DSP, le SWV ne prend pas en compte les retards impairs mais double l'auto-covariance des retards pairs¹.

Si la séquence x est *non-stationnaire et blanche*, alors $r_{xx}[t, \tau]$ est nul pour $\tau \neq 0$. On a alors :

$$s_{xx}[t, f] = 2 r_{xx}[t, 0] \quad (22)$$

$$= s_{xx}[t] \quad (23)$$

Le SWV croisé de deux séquences aléatoires centrées x et y est défini par :

$$s_{xy}[t, f] = 2 \sum_{\tau} \mathbb{E}\{x[t + \tau] \bar{y}[t - \tau]\} e^{-j4\pi f \tau} \quad (24)$$

$$= 2 \sum_{\tau} r_{xy}[t - \tau, 2\tau] e^{-j4\pi f \tau} \quad (25)$$

¹Ceci est du au parti pris de faire apparaître dans la définition du SWV l'instant central t , alors que dans la définition de la covariance t représente un instant marginal.

Le Spectre de Wigner-Ville Spatial (SWVS) de la séquence vectorielle \mathbf{x} est défini par :

$$\mathcal{S}_{\mathbf{xx}}[t, f) = 2 \sum_{\tau} \mathbb{E}\{\mathbf{x}[t + \tau] \mathbf{x}[t - \tau]^H\} e^{-j4\pi f \tau} \quad (26)$$

$$= 2 \sum_{\tau} \mathcal{R}_{\mathbf{xx}}[t - \tau, 2\tau] e^{-j4\pi f \tau} \quad (27)$$

Chapitre 1

Identifiabilité des mélanges linéaires instantanés

Nous traitons dans ce chapitre un ensemble de propriétés théoriques liées à la séparation de mélanges linéaires instantanés et notamment à leur identifiabilité. Dans le paragraphe 1.1 nous introduisons le modèle linéaire instantané, accompagné de quelques hypothèses fondamentales. Nous y discutons des indéterminations sur le gain et l'ordre des sources de la SAS. Ceci nous amènera à introduire le concept d'*identifiabilité*, puis à présenter quelques modèles de sources identifiables dans le paragraphe 1.3. Auparavant nous aurons introduit au paragraphe 1.2 le concept d'*orthonormalisation* d'un mélange, nécessaire pour prouver l'identifiabilité des modèles que nous présentons, et dont nous discuterons l'intérêt en pratique. De nombreuses notions introduites dans ce chapitre sont inspirées des travaux de Tong, Liu, Soon et Huang sur l'identifiabilité des mélanges linéaires instantanés [TLSH91].

1.1 Généralités

1.1.1 Hypothèses fondamentales

Considérons le modèle linéaire instantané suivant :

$$\mathbf{x}[t] = \mathbf{A} \mathbf{s}[t] + \mathbf{n}[t] \quad (1.1)$$

où $\forall t \in \llbracket 0, T-1 \rrbracket$:

- $\mathbf{x}[t] = [x_1[t], \dots, x_m[t]]^T$ est le vecteur de dimension m contenant les observations,
- $\mathbf{s}[t] = [s_1[t], \dots, s_n[t]]^T$ est le vecteur de dimension n contenant les sources,
- \mathbf{A} est la matrice de mélange,
- $\mathbf{n}[t] = [n_1[t], \dots, n_m[t]]^T$ est un vecteur de dimension m contenant le bruit.

On rappelle qu'avec les notations définies précédemment, on note :

$$\mathbf{s} = [\mathbf{s}[0], \dots, \mathbf{s}[T-1]] \quad (1.2)$$

et

$$\forall i \in \llbracket 1, n \rrbracket, \quad s_i = [s_i[0], \dots, s_i[T-1]] \quad (1.3)$$

D'où $\mathbf{s} = [s_1^T, \dots, s_n^T]^T$, et on dira que s_1, \dots, s_n sont les composantes de \mathbf{s} .

On suppose en outre :

- (H1) $m \geq n$,
- (H2) \mathbf{A} est une matrice de dimensions $m \times n$ de rang plein, c-à-d, $\text{rang}(\mathbf{A}) = n$,

- (H3) \mathbf{s} est la réalisation d'une séquence aléatoire vectorielle centrée, de composantes non nulles,
- (H4) \mathbf{n} est la réalisation d'une séquence aléatoire vectorielle centrée i.i.d de covariance $\sigma^2 \mathbf{I}_m$, indépendante de \mathbf{s} .

Sauf indication contraire, les signaux et la matrice de mélange seront supposés être à valeurs dans \mathbb{C} .

On note $\mathbb{H}_{\mathbf{A}}$ l'ensemble des matrices vérifiant (H1) et (H2). On note $\mathbb{H}_{\mathbf{s}}$ la modélisation des sources induite par (H3), et on utilisera (abusivement) la notation $\mathbf{s} \in \mathbb{H}_{\mathbf{s}}$ pour signifier que \mathbf{s} répond aux hypothèses définissant $\mathbb{H}_{\mathbf{s}}$.

1.1.2 Indéterminations

Nous introduisons dans ce paragraphe la notion d'espace d'identification, qui contient l'ensemble des matrices et signaux $(\mathbf{A}', \mathbf{s}') \in \mathbb{H}_{\mathbf{A}} \times \mathbb{H}_{\mathbf{s}}$ satisfaisant l'équation (1.1).

Définition 1 (Espace d'identification)

Soit $\mathbf{e} \in \mathbb{C}^{m \times T}$. L'espace d'identification de \mathbf{e} , noté $\mathbb{I}_{\mathbf{e}}$, est défini par :

$$\mathbb{I}_{\mathbf{e}} = \{(\mathbf{A}', \mathbf{s}') \in \mathbb{H}_{\mathbf{A}} \times \mathbb{H}_{\mathbf{s}} \mid \mathbf{e} = \mathbf{A}' \mathbf{s}'\} \quad (1.4)$$

Concernant notre problème de SAS on a donc $(\mathbf{A}, \mathbf{s}) \in \mathbb{I}_{(\mathbf{x}-\mathbf{n})}$. En outre, $(\mathbf{A}', \mathbf{s}') \in \mathbb{I}_{(\mathbf{x}-\mathbf{n})}$ ssi $\mathbf{A}' \mathbf{s}' = \mathbf{A} \mathbf{s}$.

Pour toute matrice inversible $\mathbf{M} \in \mathbb{C}^{n \times n}$, on a $(\mathbf{A} \mathbf{M}, \mathbf{M}^{-1} \mathbf{s}) \in \mathbb{I}_{(\mathbf{x}-\mathbf{n})}$. Il apparaît donc que sans autres hypothèses sur les sources et la matrice de mélange que celles caractérisant l'ensemble $\mathbb{H}_{\mathbf{A}} \times \mathbb{H}_{\mathbf{s}}$, le problème de la séparation de sources décrit par le modèle (1.1) ne possède pas une unique solution mais qu'au contraire il possède une forte indétermination.

Cependant, il est clair que n'importe quel élément $(\mathbf{A}', \mathbf{s}')$ de l'espace d'identification $\mathbb{I}_{(\mathbf{x}-\mathbf{n})}$ n'est pas une bonne estimation de (\mathbf{A}, \mathbf{s}) dans la mesure où il ne possède pas obligatoirement les mêmes caractéristiques que (\mathbf{A}, \mathbf{s}) , telles que les *formes d'ondes* ou certaines propriétés statistiques.

Dans de nombreuses applications, seule la forme d'onde des signaux que l'on cherche à extraire importe vraiment, et notamment plus que le *gain* (c-à-d, l'énergie) des sources ou leur numéro d'*ordre* (qui est une caractéristique immatérielle). Ainsi, si $\mathbf{D} \in \mathbb{C}^{n \times n}$ est une matrice diagonale inversible et $\mathbf{P} \in \mathbb{R}^{n \times n}$ une matrice de permutation, les signaux $\mathbf{s}' = \mathbf{P} \mathbf{D} \mathbf{s}$ peuvent être considérés comme étant une estimation satisfaisante des sources. Cette notion peut se formaliser par la définition d'une relation d'équivalence sur $\mathbb{I}_{(\mathbf{x}-\mathbf{n})}$, dite de *conservation de la forme d'onde*.

Définition 2 (Relation de conservation de la forme d'onde)

Soient $(\mathbf{A}_1, \mathbf{s}_1)$ et $(\mathbf{A}_2, \mathbf{s}_2)$ deux éléments de $\mathbb{H}_{\mathbf{A}} \times \mathbb{H}_{\mathbf{s}}$. $(\mathbf{A}_1, \mathbf{s}_1)$ et $(\mathbf{A}_2, \mathbf{s}_2)$ vérifient la *relation de conservation de la forme d'onde* suivante, notée $(\mathbf{A}_1, \mathbf{s}_1) \sim (\mathbf{A}_2, \mathbf{s}_2)$ si l'on a :

$$\mathbf{s}_1 = \mathbf{P} \mathbf{D} \mathbf{s}_2 \quad (1.5)$$

$$\mathbf{A}_1 = \mathbf{A}_2 \mathbf{D}^{-1} \mathbf{P}^T \quad (1.6)$$

où $\mathbf{D} \in \mathbb{C}^{n \times n}$ est diagonale inversible et $\mathbf{P} \in \mathbb{R}^{n \times n}$ est une matrice de permutation.

Lorsque cette notation ne présente pas d'ambiguïté, on notera abusivement $\mathbf{s}_1 \sim \mathbf{s}_2$ la relation (1.5) et $\mathbf{A}_1 \sim \mathbf{A}_2$ la relation (1.6). Notons que, puisque $m \geq n$, pour $(\mathbf{A}_1, \mathbf{s}_1)$ et $(\mathbf{A}_2, \mathbf{s}_2) \in \mathbb{I}_{(\mathbf{x}-\mathbf{n})}$ (et donc $\mathbf{A}_1 \mathbf{s}_1 = \mathbf{A}_2 \mathbf{s}_2$), si $\mathbf{A}_1 \sim \mathbf{A}_2$ alors $(\mathbf{A}_1, \mathbf{s}_1) \sim (\mathbf{A}_2, \mathbf{s}_2)$. Il suffit donc de montrer $\mathbf{A}_1 \sim \mathbf{A}_2$ pour montrer $(\mathbf{A}_1, \mathbf{s}_1) \sim (\mathbf{A}_2, \mathbf{s}_2)$ ¹.

¹Ceci n'est pas vrai dans le cas $n > m$ et constitue toute la difficulté de la séparation de mélanges sous-déterminés.

Propriété 1 (Relation d'équivalence)

La relation de conservation de la forme d'onde est une relation d'équivalence (c-à-d qu'elle possède les propriétés de réflexivité, symétrie et transitivité).

La preuve de cette propriété découle directement de la définition 2. On dira donc de deux éléments de $\mathbb{H}_{\mathbf{A}} \times \mathbb{H}_{\mathbf{s}}$ vérifiant la relation de conservation de la forme d'onde qu'ils sont *équivalents*.

Propriété 2 (Propriétés statistiques de la relation d'équivalence \sim)

Soient $(\mathbf{A}_1, \mathbf{s}_1)$ et $(\mathbf{A}_2, \mathbf{s}_2)$ deux éléments de $\mathbb{H}_{\mathbf{A}} \times \mathbb{H}_{\mathbf{s}}$ tels que $(\mathbf{A}_1, \mathbf{s}_1) \sim (\mathbf{A}_2, \mathbf{s}_2)$. Soient $s_{1,1}, \dots, s_{1,n}$ et $s_{2,1}, \dots, s_{2,n}$ les composantes de \mathbf{s}_1 et \mathbf{s}_2 . Alors on a les propriétés suivantes :

- a) Les séquences $s_{1,1}, \dots, s_{1,n}$ sont mutuellement indépendantes ssi les séquences $s_{2,1}, \dots, s_{2,n}$ sont mutuellement indépendantes.
- b) $\mathcal{R}_{\mathbf{s}_1 \mathbf{s}_1}[t, \tau]$ est diagonale ssi $\mathcal{R}_{\mathbf{s}_2 \mathbf{s}_2}[t, \tau]$ est diagonale.

Preuve de la propriété 2

- a) D'après (1.5), $\forall i \in \llbracket 1, n \rrbracket$, $s_{2,i} = \lambda_i s_{1,p(i)}$, où p désigne une permutation de $\llbracket 1, n \rrbracket$ et où $\lambda_i \in \mathbb{C}, \lambda_i \neq 0$. Donc les séquences $\{s_{2,1}, \dots, s_{2,n}\}$ sont mutuellement indépendantes ssi les séquences $\{s_{1,1}, \dots, s_{1,n}\}$ sont mutuellement indépendantes.
- b) D'après (1.5), on a $\mathcal{R}_{\mathbf{s}_2 \mathbf{s}_2}[t, \tau] = \mathbf{P} \mathbf{D} \mathcal{R}_{\mathbf{s}_1 \mathbf{s}_1}[t, \tau] \mathbf{D}^H \mathbf{P}^T$. Si $\mathcal{R}_{\mathbf{s}_1 \mathbf{s}_1}[t, \tau]$ est diagonale, alors $\mathcal{R}_{\mathbf{s}_2 \mathbf{s}_2}[t, \tau]$ est diagonale, et inversement. \square

L'intérêt de la relation de conservation de la forme d'onde \sim réside dans le fait que, puisque c'est une relation d'équivalence sur $\mathbb{H}_{\mathbf{A}} \times \mathbb{H}_{\mathbf{s}}$, elle définit une partition de $\mathbb{H}_{\mathbf{A}} \times \mathbb{H}_{\mathbf{s}}$. Ainsi, le but de la séparation aveugle du mélange défini par (1.1) sera d'identifier la classe d'équivalence à laquelle appartient (\mathbf{A}, \mathbf{s}) , ce qui revient à estimer n'importe quel élément de cette classe.

Nous avons vu précédemment que sans autres hypothèses sur les sources et la matrice de mélange que celles définissant l'ensemble $\mathbb{H}_{\mathbf{A}} \times \mathbb{H}_{\mathbf{s}}$ (paragraphe 1.1.1), il n'apparaît pas être possible d'identifier exactement (\mathbf{A}, \mathbf{s}) , ni même la classe d'équivalence à laquelle appartient (\mathbf{A}, \mathbf{s}) . En revanche nous allons montrer qu'en ajoutant des contraintes sur les sources, le problème peut devenir *identifiable*.

1.1.3 Identifiabilité

Nous introduisons dans ce paragraphe la notion d'identifiabilité.

Définition 3 (Identifiabilité)

Supposons que les sources \mathbf{s} satisfont une série d'hypothèses (supplémentaires à $\mathbb{H}_{\mathbf{s}}$) qui définissent un modèle² que l'on note \mathbb{M} . Alors on dit que la matrice de mélange et les sources $(\mathbf{A}, \mathbf{s}) \in \mathbb{I}_{(\mathbf{x}-\mathbf{n})}$ sont *identifiables selon \mathbb{M}* ssi $\forall (\mathbf{A}', \mathbf{s}') \in \mathbb{I}_{(\mathbf{x}-\mathbf{n})}$:

$$\mathbf{s}' \in \mathbb{M} \Rightarrow (\mathbf{A}', \mathbf{s}') \sim (\mathbf{A}, \mathbf{s})$$

En d'autres termes, un modèle \mathbb{M} des sources est identifiable s'il suffit de trouver $(\mathbf{A}', \mathbf{s}')$ tel que $\mathbf{x} = \mathbf{A}' \mathbf{s}' + \mathbf{n}$ et tel que $\mathbf{s}' \in \mathbb{M}$ pour que \mathbf{s}' soit une estimation de \mathbf{s} à permutations et gains près. Il est important de souligner que l'identifiabilité est définie relativement à un modèle.

Remarquons que (\mathbf{A}, \mathbf{s}) n'est pas identifiable selon $\mathbb{H}_{\mathbf{s}}$ car $\forall \mathbf{M} \in \mathbb{C}^{n \times n}$ inversible et non diagonale, on a $(\mathbf{A} \mathbf{M}, \mathbf{M}^{-1} \mathbf{s}) \in \mathbb{I}_{(\mathbf{x}-\mathbf{n})}$ et $\mathbf{M}^{-1} \mathbf{s} \in \mathbb{H}_{\mathbf{s}}$ mais $(\mathbf{A} \mathbf{M}, \mathbf{M}^{-1} \mathbf{s}) \not\sim (\mathbf{A}, \mathbf{s})$. D'où la nécessité d'ajouter des hypothèses sur \mathbf{s} pour rendre le problème identifiable.

Nous présentons au paragraphe 1.3 plusieurs modèles identifiables mais afin de prouver l'identifiabilité de ces modèles nous devons introduire auparavant le concept d'*orthonormalisation*.

²A nouveau, on utilisera abusivement la notation $\mathbf{s} \in \mathbb{M}$ pour signifier que \mathbf{s} répond aux hypothèses définissant le modèle \mathbb{M} .

1.2 Orthonormalisation

Le concept d'orthonormalisation, largement utilisé en SAS, est présenté avec rigueur dans [TLSH91]. Nous en proposons une formulation plus simple et à notre avis plus claire, basée sur la propriété suivante.

Propriété 3 (Orthonormalisation)

Soit $\mathbf{W} \in \mathbb{C}^{n \times m}$ une matrice de rang plein vérifiant :

$$\mathbf{W} \mathbf{A} \mathbf{M} \mathbf{A}^H \mathbf{W}^H = \mathbf{I}_n \quad (1.7)$$

où \mathbf{M} est une matrice hermitienne définie positive et $\mathbf{A} \in \mathbb{H}_{\mathbf{A}}$. On pose :

$$\mathbf{U} = \mathbf{W} \mathbf{A} \mathbf{M}^{\frac{1}{2}} \quad (1.8)$$

Alors on a :

$$\mathbf{U}^H \mathbf{W} = \mathbf{M}^{-\frac{1}{2}} \mathbf{A}^\# \quad (1.9)$$

et

$$\mathbf{W}^\# \mathbf{U} = \mathbf{A} \mathbf{M}^{\frac{1}{2}} \quad (1.10)$$

Cette propriété montre qu'il est possible, connaissant $\mathbf{A} \mathbf{M} \mathbf{A}^H$, de calculer la matrice $\mathbf{A} \mathbf{M}^{\frac{1}{2}}$ en deux étapes : en calculant dans un premier temps une matrice \mathbf{W} vérifiant la propriété (1.7) puis en calculant la matrice $\mathbf{U} = \mathbf{W} \mathbf{A} \mathbf{M}^{\frac{1}{2}}$. L'intérêt de cette décomposition est que la matrice \mathbf{U} est orthonormale (d'où le nom d'«orthonormalisation») et nous verrons dans les prochains paragraphes que cette propriété est très appréciable dans l'élaboration de méthodes de séparation. Nous verrons aussi que dans le cas des mélanges linéaires instantanés la matrice \mathbf{M} modélise l'indétermination sur le gain de la SAS. En effet, en pratique, bien que \mathbf{A} soit inconnue, nous serons en mesure d'estimer dans un premier temps une quantité de la forme $\mathbf{A} \mathbf{M} \mathbf{A}^H$, avec \mathbf{M} inconnue mais diagonale, puis d'estimer $\mathbf{A} \mathbf{M}^{\frac{1}{2}}$, équivalente à \mathbf{A} .

Dans le cas des mélanges convolutifs, qui seront présentés dans le chapitre 4, nous verrons que \mathbf{M} modélise une indétermination à un filtrage près sur les sources. En effet, il sera montré qu'un mélange convolutif RIF peut être ramené à un mélange instantané. Nous serons en mesure d'estimer une quantité de la forme $\mathbf{A} \mathbf{M} \mathbf{A}^H$, avec cette fois \mathbf{M} *bloc*-diagonale, puis d'estimer $\mathbf{A} \mathbf{M}^{\frac{1}{2}}$, équivalente à \mathbf{A} à une matrice *bloc*-diagonale près.

Avant de démontrer la propriété 3, étudions l'existence de \mathbf{W} . \mathbf{A} étant une matrice de rang n et \mathbf{M} étant hermitienne inversible, $\mathbf{A} \mathbf{M} \mathbf{A}^H$ est une matrice hermitienne de dimensions $m \times m$ et de rang n . Elle possède donc $m - n$ valeurs propres nulles. Soient $\lambda_1, \dots, \lambda_n$ ses valeurs propres strictement positives et $\mathbf{h}_1, \dots, \mathbf{h}_n$ des vecteurs propres normés correspondants. On a $\forall (i, j) \in \llbracket 1, n \rrbracket^2, i \neq j$:

$$\mathbf{h}_i^H \mathbf{A} \mathbf{M} \mathbf{A}^H \mathbf{h}_i = \lambda_i \quad (1.11)$$

$$\mathbf{h}_j^H \mathbf{A} \mathbf{M} \mathbf{A}^H \mathbf{h}_i = 0 \quad (1.12)$$

Soit \mathbf{W} la matrice $\in \mathbb{C}^{n \times m}$ définie par :

$$\mathbf{W} = \left[\frac{1}{\sqrt{\lambda_1}} \mathbf{h}_1 \dots \frac{1}{\sqrt{\lambda_n}} \mathbf{h}_n \right]^H \quad (1.13)$$

Par construction, on a :

$$\mathbf{W} \mathbf{A} \mathbf{M} \mathbf{A}^H \mathbf{W}^H = \mathbf{I}_m \quad (1.14)$$

Par construction de rang plein et satisfaisant l'équation (1.7), \mathbf{W} vérifie donc les hypothèses de la propriété 3 - sans être toutefois l'unique matrice vérifiant ces hypothèses. On remarquera qu'ainsi construite, les lignes de \mathbf{W} sont orthogonales :

$$\mathbf{W} \mathbf{W}^H = \text{diag} \left(\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n} \right) \quad (1.15)$$

Montrons à présent la propriété 3.

Preuve de la propriété 3

Soit $\tilde{\mathbf{A}} = \mathbf{A} \mathbf{M}^{\frac{1}{2}}$. On a $\mathbf{U} = \mathbf{W} \tilde{\mathbf{A}}$. Montrons dans un premier temps $\mathbf{W}^\# \mathbf{U} = \tilde{\mathbf{A}}$. Soit la décomposition en valeurs singulières suivante de $\tilde{\mathbf{A}}$:

$$\tilde{\mathbf{A}} = \mathbf{U}_1 \mathbf{\Sigma} \mathbf{U}_2 \quad (1.16)$$

où :

- $\mathbf{U}_1 \in \mathbb{C}^{m \times n}$ vérifie $\mathbf{U}_1^H \mathbf{U}_1 = \mathbf{I}_n$,
- $\mathbf{\Sigma} \in \mathbb{R}^{n \times n}$ est diagonale > 0 ,
- $\mathbf{U}_2 \in \mathbb{C}^{n \times n}$ est orthonormale.

On a alors $\tilde{\mathbf{A}} \tilde{\mathbf{A}}^H = \mathbf{U}_1 \mathbf{\Sigma}^2 \mathbf{U}_1^H$, soit encore :

$$\mathbf{U}_1^H (\tilde{\mathbf{A}} \tilde{\mathbf{A}}^H) \mathbf{U}_1 = \mathbf{\Sigma}^2 \quad (1.17)$$

Soit $\mathbf{N} = \mathbf{W} \mathbf{W}^H$. On pose :

$$\mathbf{V} = \mathbf{W}^H \mathbf{N}^{-\frac{H}{2}} \in \mathbb{C}^{m \times n} \quad (1.18)$$

On a :

$$\mathbf{V}^H \mathbf{V} = \mathbf{N}^{-\frac{1}{2}} \mathbf{W} \mathbf{W}^H \mathbf{N}^{-\frac{H}{2}} \quad (1.19)$$

$$= \mathbf{N}^{-\frac{1}{2}} \mathbf{N} \mathbf{N}^{-\frac{H}{2}} \quad (1.20)$$

$$= \mathbf{N}^{-\frac{1}{2}} \mathbf{N}^{\frac{1}{2}} \mathbf{N}^{\frac{H}{2}} \mathbf{N}^{-\frac{H}{2}} \quad (1.21)$$

$$= \mathbf{I}_n \quad (1.22)$$

Les colonnes de \mathbf{V} sont donc orthonormales. En outre, on a :

$$\mathbf{V}^H (\tilde{\mathbf{A}} \tilde{\mathbf{A}}^H) \mathbf{V} = \mathbf{N}^{-\frac{1}{2}} \mathbf{W} (\tilde{\mathbf{A}} \tilde{\mathbf{A}}^H) \mathbf{W}^H \mathbf{N}^{-\frac{H}{2}} \quad (1.23)$$

$$= \mathbf{N}^{-1} \quad (1.24)$$

\mathbf{N} est par construction une matrice hermitienne définie positive. Soit $\mathbf{N} = \mathbf{L} \mathbf{D} \mathbf{L}^H$ la décomposition en valeurs propres de \mathbf{N} . Il vient donc :

$$(\mathbf{V} \mathbf{L})^H (\tilde{\mathbf{A}} \tilde{\mathbf{A}}^H) (\mathbf{V} \mathbf{L}) = \mathbf{D}^{-1} \quad (1.25)$$

On en déduit donc d'après (1.17), $\mathbf{D}^{-1} = \mathbf{\Sigma}^2$ (éventuellement à permutation de termes diagonaux près) et $\mathbf{V} \mathbf{L}$ est égale à \mathbf{U}_1 (à permutations de colonnes près, voire à "rotations" de colonnes près si \mathbf{A} possède des valeurs singulières de multiplicité supérieure à 1). Il existe donc $\mathbf{Q} \in \mathbb{C}^{n \times n}$ orthonormale telle que :

$$\tilde{\mathbf{A}} = (\mathbf{V} \mathbf{L}) \mathbf{\Sigma} \mathbf{Q} \quad (1.26)$$

\mathbf{Q} peut s'écrire :

$$\mathbf{Q} = \mathbf{\Sigma}^{-1} (\mathbf{V} \mathbf{L})^H \tilde{\mathbf{A}} \quad (1.27)$$

D'où, avec $\Sigma^{-1} = \mathbf{D}^{\frac{1}{2}}$ et $\mathbf{N}^{-\frac{1}{2}} = \mathbf{L} \mathbf{D}^{-\frac{1}{2}} \mathbf{L}^H$,

$$\mathbf{Q} = \Sigma^{-1} \mathbf{L}^H \mathbf{N}^{-\frac{1}{2}} (\mathbf{W} \tilde{\mathbf{A}}) \quad (1.28)$$

$$= \mathbf{D}^{\frac{1}{2}} \mathbf{L}^H (\mathbf{L} \mathbf{D}^{-\frac{1}{2}} \mathbf{L}^H) \mathbf{U} \quad (1.29)$$

$$= \mathbf{L}^H \mathbf{U} \quad (1.30)$$

D'où :

$$\tilde{\mathbf{A}} = \mathbf{V} (\mathbf{L} \Sigma \mathbf{L}^H) \mathbf{U} \quad (1.31)$$

$$= \mathbf{V} \mathbf{N}^{-\frac{1}{2}} \mathbf{U} \quad (1.32)$$

$$= \mathbf{W}^H (\mathbf{W} \mathbf{W}^H)^{-1} \mathbf{U} \quad (1.33)$$

$$= \mathbf{W}^\# \mathbf{U} \quad (1.34)$$

Finalement :

$$\mathbf{A} \mathbf{M}^{\frac{1}{2}} = \mathbf{W}^\# \mathbf{U} \quad (1.35)$$

En outre la pseudo-inverse de $\tilde{\mathbf{A}}$ s'écrit :

$$\tilde{\mathbf{A}}^\# = \mathbf{Q}^H \Sigma^{-1} \mathbf{L}^H \mathbf{V}^H \quad (1.36)$$

$$= \mathbf{U}^H (\mathbf{L} \Sigma^{-1} \mathbf{L}^H) \mathbf{V}^H \quad (1.37)$$

$$= \mathbf{U}^H \mathbf{N}^{\frac{1}{2}} \mathbf{N}^{-\frac{1}{2}} \mathbf{W} \quad (1.38)$$

$$= \mathbf{U}^H \mathbf{W} \quad (1.39)$$

Finalement :

$$\mathbf{M}^{-\frac{1}{2}} \mathbf{A}^\# = \mathbf{U}^H \mathbf{W} \quad (1.40)$$

□

1.3 Modèles identifiables

Nous présentons dans cette partie quelques modèles de sources identifiables. Ces modèles possèdent une hypothèse commune : *les sources sont modélisées par des séquences aléatoires mutuellement décorréliées*. Selon la nature des sources, cette hypothèse peut être suffisante ou non et deux types de modèles sont à envisager :

- si les sources possèdent une *diversité temporelle et/ou fréquentielle*, c-à-d si leur SWV varie en temps et/ou en fréquence alors la décorrélation mutuelle des sources suffit à leur identifiabilité. Le premier modèle \mathbb{M}_1 que nous présenterons dans le paragraphe 1.3.1 considère le cas général de sources non-stationnaires colorées pour lequel le SWV des sources dépend effectivement du temps *et* de la fréquence. Puis nous présenterons deux cas particuliers : le modèle \mathbb{M}_2 qui considère le cas de sources non-stationnaires blanches (c-à-d que le SWV des sources est constant en fréquence) puis le modèle \mathbb{M}_3 qui considère le cas de sources stationnaires colorées (c-à-d que le SWV des sources est constant dans le temps).
- si les sources ne possèdent pas de *diversité temporelle ou fréquentielle*, c-à-d si leur SWV est constant en temps et en fréquence alors la décorrélation mutuelle des sources ne suffit plus à leur identifiabilité. Ce cas de figure correspond aux hypothèses du modèle \mathbb{M}_4 présenté au paragraphe 1.3.4 dans lequel les sources sont modélisées par des séquences aléatoires i.i.d. Nous montrerons que l'identifiabilité de ce modèle est garantie si les sources sont mutuellement indépendantes (au sens strict) et à condition qu'au plus une source soit de distribution gaussienne.

1.3.1 Sources non-stationnaires colorées et mutuellement décorréées (modèle \mathbb{M}_1)

Supposons que les sources possèdent une *diversité temporelle et/ou fréquentielle*, et qu'on les modélise par des séquences aléatoires non-stationnaires et colorées.

Les statistiques des sources varient alors au cours du temps, en particulier, leur covariance $\mathcal{R}_{ss}[t, \tau]$ dépend du temps t et du retard τ , ou de manière équivalente, *leur SWVS $\mathcal{S}_{ss}[t, f]$ dépend du temps t et de la fréquence f .*

On pose :

$$\mathcal{R}_{ss} = \frac{1}{T} \sum_{t=0}^{T-1} \mathcal{R}_{ss}[t, 0] \quad (1.41)$$

qui définit la «puissance temporelle moyenne» des sources.

Propriété 4

Soit \mathbb{M}_1 le modèle défini par :

- (M1) Les sources sont des séquences non-stationnaires colorées,
- (M2) Les sources sont décorréées :

$$\Rightarrow \quad \forall(t, f), \quad \mathcal{S}_{ss}[t, f] = \text{diag}(s_{s_1 s_1}[t, f], \dots, s_{s_n s_n}[t, f])$$

(En particulier : $\mathcal{R}_{ss} = \text{diag}(r_{s_1 s_1}, \dots, r_{s_n s_n})$)

(M3) $\exists (t_1, f_1)$ tel que

$$\forall i \neq j, \quad \frac{s_{s_i s_i}[t_1, f_1]}{r_{s_i s_i}} \neq \frac{s_{s_j s_j}[t_1, f_1]}{r_{s_j s_j}}$$

Le modèle \mathbb{M}_1 est identifiable.

Preuve de la propriété 4

Soit $(\mathbf{A}', \mathbf{s}') \in \mathbb{I}_{(\mathbf{x}-\mathbf{n})}$ tel que $\mathbf{s}' \in \mathbb{M}_1$. Montrons $(\mathbf{A}', \mathbf{s}') \sim (\mathbf{A}, \mathbf{s})$. Nous avons vu pour $m \geq n$ qu'il suffit de montrer $\mathbf{A}' \sim \mathbf{A}$.

Soit $\mathbf{W} \in \mathbb{C}^{n \times m}$ de rang plein vérifiant :

$$\mathbf{W} \mathbf{A} \mathcal{R}_{ss} \mathbf{A}^H \mathbf{W}^H = \mathbf{I}_n \quad (1.42)$$

Avec $\mathbf{A} \mathbf{s} = \mathbf{A}' \mathbf{s}'$, on a aussi :

$$\mathbf{W} \mathbf{A}' \mathcal{R}_{s' s'} \mathbf{A}'^H \mathbf{W}^H = \mathbf{I}_n \quad (1.43)$$

En définissant $\mathbf{U} = \mathbf{W} \mathbf{A} \mathcal{R}_{ss}^{\frac{1}{2}}$ et $\mathbf{U}' = \mathbf{W} \mathbf{A}' \mathcal{R}_{s' s'}^{\frac{1}{2}}$, on a donc, d'après la propriété 3 :

$$\mathbf{A} \mathcal{R}_{ss}^{\frac{1}{2}} = \mathbf{W}^\# \mathbf{U} \quad (1.44)$$

$$\mathbf{A}' \mathcal{R}_{s' s'}^{\frac{1}{2}} = \mathbf{W}^\# \mathbf{U}' \quad (1.45)$$

Montrons $\mathbf{U} \sim \mathbf{U}'$.

Soit $\mathcal{R} = \mathbf{W} \mathbf{A} \mathcal{S}_{ss}[t_1, f_1] \mathbf{A}^H \mathbf{W}^H = \mathbf{W} \mathbf{A}' \mathcal{S}_{s' s'}[t_1, f_1] \mathbf{A}'^H \mathbf{W}^H$. On a :

$$\mathcal{R} = \mathbf{U} \mathcal{R}_{ss}^{-\frac{1}{2}} \mathcal{S}_{ss}[t_1, f_1] \mathcal{R}_{ss}^{-\frac{H}{2}} \mathbf{U}^H \quad (1.46)$$

$$= \mathbf{U} \mathcal{R}_{ss}^{-1} \mathcal{S}_{ss}[t_1, f_1] \mathbf{U}^H \quad (1.47)$$

De la même façon :

$$\mathcal{R} = \mathbf{U}' \mathcal{R}_{s' s'}^{-\frac{1}{2}} \mathcal{S}_{s' s'}[t_1, f_1] \mathcal{R}_{s' s'}^{-\frac{H}{2}} \mathbf{U}'^H \quad (1.48)$$

$$= \mathbf{U}' \mathcal{R}_{s' s'}^{-1} \mathcal{S}_{s' s'}[t_1, f_1] \mathbf{U}'^H \quad (1.49)$$

\mathbf{U} et \mathbf{U}' sont donc deux bases orthonormales dans lesquelles \mathcal{R} est diagonale. Par hypothèse les coefficients diagonaux de $\mathcal{R}_{\mathbf{ss}}^{-1} \mathcal{S}_{\mathbf{ss}}[t_1, f_1]$ et $\mathcal{R}_{\mathbf{s's'}}^{-1} \mathcal{S}_{\mathbf{s's'}}[t_1, f_1]$ sont distincts. On en déduit :

$$\mathbf{U}' = \mathbf{U} \mathbf{P} \mathbf{\Phi} \quad (1.50)$$

où $\mathbf{P} \in \mathbb{R}^{n \times n}$ est une matrice de permutation et $\mathbf{\Phi} \in \mathbb{C}^{n \times n}$ est une matrice diagonale à coefficients de module unitaire. D'où :

$$\mathbf{A}' \mathcal{R}_{\mathbf{s's'}}^{\frac{1}{2}} = \mathbf{W}^\# \mathbf{U} \mathbf{P} \mathbf{\Phi} \quad (1.51)$$

$$= \mathbf{A} \mathcal{R}_{\mathbf{ss}}^{\frac{1}{2}} \mathbf{P} \mathbf{\Phi} \quad (1.52)$$

et donc :

$$\mathbf{A}' = \mathbf{A} \mathcal{R}_{\mathbf{ss}}^{\frac{1}{2}} \mathbf{P} \mathbf{\Phi} \mathcal{R}_{\mathbf{s's'}}^{-\frac{1}{2}} \quad (1.53)$$

c-à-d, $\mathbf{A}' \sim \mathbf{A}$. \square

On peut constater dès à présent que la preuve de la propriété 4 est «constructive». En effet, en supposant pour simplifier le bruit nul ($\mathbf{n} = \mathbf{0}$), si l'on peut estimer en pratique \mathbf{W} et $\mathcal{S}_{\mathbf{xx}}[t_1, f_1] = \mathbf{A} \mathcal{S}_{\mathbf{ss}}[t_1, f_1] \mathbf{A}^H$, on voit que pour estimer la matrice $\mathbf{A} \mathcal{R}_{\mathbf{ss}}^{\frac{1}{2}}$ équivalente à \mathbf{A} , il suffit de diagonaliser $\mathcal{R} = \mathbf{W} \mathcal{S}_{\mathbf{xx}}[t_1, f_1] \mathbf{W}^H$ afin d'obtenir \mathbf{U} .

Nous verrons par la suite que le principal enjeu des méthodes de séparation que nous étudierons s'avère être l'estimation de la quantité $\mathcal{S}_{\mathbf{xx}}[t, f]$.

On peut montrer que la condition (M3) de \mathbb{M}_1 peut être assouplie. En effet, plutôt que de supposer qu'il existe un point (t_1, f_1) fixé tel que les termes diagonaux de $\mathcal{R}_{\mathbf{ss}}^{-1} \mathcal{S}_{\mathbf{ss}}[t_1, f_1]$ soient tous distincts (afin de garantir dans la preuve que \mathbf{U} et \mathbf{U}' sont équivalentes), il suffit de supposer l'existence d'un ensemble de points $\{(t_1, f_1), \dots, (t_K, f_K)\}$ tels que pour chaque paire de termes diagonaux aux positions i et j , il existe un point (t_k, f_k) pour lequel les deux termes diagonaux aux positions i et j de $\mathcal{R}_{\mathbf{ss}}^{-1} \mathcal{S}_{\mathbf{ss}}[t_k, f_k]$ soient distincts. La preuve mathématique de cette affirmation repose sur l'unicité de la diagonalisation simultanée, notion qui sera présentée ultérieurement. La preuve est donnée dans [BAMCM97]. Ceci nous amène au modèle \mathbb{M}'_1 , moins contraint que \mathbb{M}_1 .

Propriété 5

Soit \mathbb{M}'_1 le modèle défini par :

(M1) Les sources sont des séquences non-stationnaires colorées,

(M2) Les sources sont décorrélées :

$$\Rightarrow \quad \forall(t, f), \quad \mathcal{S}_{\mathbf{ss}}[t, f] = \text{diag}(s_{s_1 s_1}[t, f], \dots, s_{s_n s_n}[t, f])$$

(M3) $\exists \{(t_1, f_1), \dots, (t_K, f_K)\}$ tels que

$$\forall i \neq j, \quad \exists k \in \llbracket 1, K \rrbracket \text{ tel que } \frac{s_{s_i s_i}[t_k, f_k]}{r_{s_i s_i}} \neq \frac{s_{s_j s_j}[t_k, f_k]}{r_{s_j s_j}}$$

Le modèle \mathbb{M}'_1 est identifiable.

Deux modèles de sources standard s'avèrent être des cas particuliers du modèle général \mathbb{M}_1 : le modèle de sources stationnaires et colorées et le modèle de sources non-stationnaires blanches.

1.3.2 Sources non-stationnaires blanches et mutuellement décorréées (modèle \mathbb{M}_2)

Si les sources sont non-stationnaires et blanches alors $\mathcal{R}_{ss}[t, \tau]$ est nul pour $\tau \neq 0$, et donc :

$$\mathcal{S}_{ss}[t, f] = 2 \mathcal{R}_{ss}[t, 0] \quad (1.54)$$

$$= \mathcal{S}_{ss}[t] \quad (1.55)$$

Le SWVS des sources $\mathcal{S}_{ss}[t, f]$ est constant en fréquence et ne dépend que du temps.

L'identifiabilité de ce modèle peut alors aussi plus simplement s'écrire en termes de covariances.

Propriété 6

Soit \mathbb{M}_2 le modèle défini par :

(M1) Les sources sont des séquences non-stationnaires blanches :

$$\forall \tau \neq 0, \quad \mathcal{R}_{ss}[t, \tau] = \mathbf{0}_n$$

(M2) Les sources sont décorréées :

$$\forall t, \quad \mathcal{R}_{ss}[t, 0] = \text{diag}(r_{s_1 s_1}[t, 0], \dots, r_{s_n s_n}[t, 0])$$

(M3) $\exists t_1$ tel que

$$\forall i \neq j, \quad \frac{r_{s_i s_i}[t_1, 0]}{r_{s_i s_i}} \neq \frac{r_{s_j s_j}[t_1, 0]}{r_{s_j s_j}}$$

Le modèle \mathbb{M}_2 est identifiable.

1.3.3 Sources stationnaires colorées et mutuellement décorréées (modèle \mathbb{M}_3)

Si les sources sont stationnaires et colorées alors on a $\mathcal{R}_{ss}[t, \tau] = \mathcal{R}_{ss}[\tau]$. En particulier, la variance des sources est constante dans le temps d'où, avec l'équation (1.41) :

$$\mathcal{R}_{ss} = \mathcal{R}_{ss}[0] \quad (1.56)$$

En outre, avec la définition du SWVS donné par l'équation (26) on a :

$$\mathcal{S}_{ss}[t, f] = 2 \sum_{\tau \text{ pair} \in \mathbb{Z}} \mathcal{R}_{ss}[\tau] e^{-j 2 \pi f \tau} \quad (1.57)$$

$$= \mathcal{S}_{ss}[f] \quad (1.58)$$

Le SWVS des sources $\mathcal{S}_{ss}[t, f]$ est constant dans le temps et ne dépend que de la fréquence.

L'identifiabilité de ce modèle peut plus simplement s'écrire en termes de covariances.

Propriété 7

Soit \mathbb{M}_3 le modèle défini par :

(M1) Les sources sont des séquences stationnaires colorées :

$$\forall (t, \tau), \quad \mathcal{R}_{ss}[t, \tau] = \mathcal{R}_{ss}[\tau]$$

(M2) Les sources sont décorréées :

$$\forall \tau, \quad \mathcal{R}_{ss}[\tau] = \text{diag}(r_{s_1 s_1}[\tau], \dots, r_{s_n s_n}[\tau])$$

(M3) $\exists \tau_1$ tel que

$$\forall i \neq j, \quad \frac{r_{s_i s_i}[\tau_1]}{r_{s_i s_i}[0]} \neq \frac{r_{s_j s_j}[\tau_1]}{r_{s_j s_j}[0]}$$

Le modèle \mathbb{M}_3 est identifiable.

Il va de soit que l'hypothèse (M3) des modèles \mathbb{M}_2 et \mathbb{M}_3 peut être assouplie de la même façon que pour \mathbb{M}_1 , donnant alors les modèles \mathbb{M}'_2 et \mathbb{M}'_3 .

1.3.4 Sources i.i.d et mutuellement indépendantes (modèle \mathbb{M}_4)

On considère dans cette partie le cas où les sources sont modélisées par des séquences aléatoires i.i.d. Le fait de modéliser des sources par des séquences i.i.d peut relever de deux motivations : 1) ou bien les sources ne possèdent pas de *diversité temporelle et/ou fréquentielle* et ceci justifie quelle soient supposées i.i.d, 2) ou bien les sources possèdent une *diversité temporelle et/ou fréquentielle* mais on choisira de ne pas l'utiliser dans la conception d'un algorithme de séparation, ce qui revient à les supposer i.i.d.

Dans ce cas, en notant $\mathcal{R}_{ss}[t] = \mathcal{R}_{ss}$ la covariance (constante) des sources, le SWVS des sources vaut simplement :

$$\mathcal{S}_{ss}[t, f] = 2 \mathcal{R}_{ss} \quad (1.59)$$

Le SWVS des sources $\mathcal{S}_{ss}[t, f]$ est constant en temps et en fréquence.

Ceci implique que l'indépendance à l'ordre deux des sources ne suffit pas à garantir leur identifiabilité. Il est alors nécessaire d'utiliser une *information discriminante* aux ordres supérieurs par exemple en supposant les sources mutuellement indépendantes au sens strict.

Propriété 8

Soit \mathbb{M}_4 le modèle défini par :

(M1) Les sources sont des séquences identiquement et indépendamment distribuées :

$$\forall i \in \llbracket 1, n \rrbracket, \quad p(s_i) = \prod_{t=0}^{T-1} p(s_i[t]) \quad \text{et} \quad \forall (t, t'), \quad p(s_i[t]) = p(s_i[t'])$$

(M2) Les sources sont mutuellement indépendantes, c-à-d

$$p(\mathbf{s}) = \prod_{i=1}^n p(s_i)$$

(M3) Une source au plus est gaussienne.

Le modèle \mathbb{M}_4 est identifiable.

La preuve de la propriété 8 repose sur un théorème de Darmois, adapté par Comon [Com94] :

Théorème 1

Soit $\mathbf{v} = [v_1, \dots, v_n]^T$ un vecteur aléatoire de dimension n , de composantes mutuellement indépendantes (c-à-d, $p(\mathbf{v}) = \prod_{i=1}^n p(v_i)$) et dont une composante au plus est gaussienne. Soit $\mathbf{C} \in \mathbb{C}^{n \times n}$ une matrice orthonormale et $\mathbf{w} = \mathbf{C} \mathbf{v}$.

Alors les composantes de \mathbf{w} sont mutuellement indépendantes ssi $\mathbf{C} = \mathbf{P} \mathbf{\Phi}$, où $\mathbf{P} \in \mathbb{R}^{n \times n}$ est une matrice de permutation et $\mathbf{\Phi} \in \mathbb{C}^{n \times n}$ est une matrice diagonale à coefficients de module unitaire.

Preuve de la propriété 8

Soit $(\mathbf{A}', \mathbf{s}') \in \mathbb{I}_{(\mathbf{x}-\mathbf{n})}$ tel que $\mathbf{s}' \in \mathbb{M}_4$. Montrons $\mathbf{A}' \sim \mathbf{A}$.

Soit $\mathcal{R}_{\mathbf{s}\mathbf{s}} = \mathcal{R}_{\mathbf{s}\mathbf{s}}[t, 0]$ et $\mathcal{R}_{\mathbf{s}'\mathbf{s}'} = \mathcal{R}_{\mathbf{s}'\mathbf{s}'}[t, 0]$ les covariances respectives de \mathbf{s} et \mathbf{s}' (constantes puisque les signaux sont i.i.d). Soit $\mathbf{W} \in \mathbb{C}^{n \times m}$ de rang plein telle que :

$$\mathbf{W} \mathbf{A} \mathcal{R}_{\mathbf{s}\mathbf{s}} \mathbf{A}^H \mathbf{W}^H = \mathbf{I}_n \quad (1.60)$$

Avec $\mathbf{A} \mathbf{s} = \mathbf{A}' \mathbf{s}'$, on a :

$$\mathbf{W} \mathbf{A}' \mathcal{R}_{\mathbf{s}'\mathbf{s}'} \mathbf{A}'^H \mathbf{W}^H = \mathbf{I}_n \quad (1.61)$$

En définissant $\mathbf{U} = \mathbf{W} \mathbf{A} \mathcal{R}_{\mathbf{s}\mathbf{s}}^{-\frac{1}{2}}$ et $\mathbf{U}' = \mathbf{W} \mathbf{A}' \mathcal{R}_{\mathbf{s}'\mathbf{s}'}^{-\frac{1}{2}}$, on a donc, d'après la propriété 3 :

$$\mathbf{A} \mathcal{R}_{\mathbf{s}\mathbf{s}}^{-\frac{1}{2}} = \mathbf{W}^\# \mathbf{U} \quad (1.62)$$

$$\mathbf{A}' \mathcal{R}_{\mathbf{s}'\mathbf{s}'}^{-\frac{1}{2}} = \mathbf{W}^\# \mathbf{U}' \quad (1.63)$$

Montrons $\mathbf{U} \sim \mathbf{U}'$.

On a $\mathbf{W} \mathbf{A} \mathbf{s} = \mathbf{W} \mathbf{A}' \mathbf{s}'$, c-à-d, $\mathbf{U} \mathcal{R}_{\mathbf{s}\mathbf{s}}^{-\frac{1}{2}} \mathbf{s} = \mathbf{U}' \mathcal{R}_{\mathbf{s}'\mathbf{s}'}^{-\frac{1}{2}} \mathbf{s}'$, soit encore :

$$\mathcal{R}_{\mathbf{s}\mathbf{s}}^{-\frac{1}{2}} \mathbf{s} = \mathbf{U}^H \mathbf{U}' \mathcal{R}_{\mathbf{s}'\mathbf{s}'}^{-\frac{1}{2}} \mathbf{s}' \quad (1.64)$$

Les composantes de \mathbf{s} et \mathbf{s}' étant par hypothèse respectivement mutuellement indépendantes, elles sont en particulier décorréées, c-à-d que les matrices $\mathcal{R}_{\mathbf{s}\mathbf{s}}$ et $\mathcal{R}_{\mathbf{s}'\mathbf{s}'}$ sont diagonales > 0 . Les composantes de $\mathcal{R}_{\mathbf{s}\mathbf{s}}^{-\frac{1}{2}} \mathbf{s}[\cdot]$ et de $\mathcal{R}_{\mathbf{s}'\mathbf{s}'}^{-\frac{1}{2}} \mathbf{s}'[\cdot]$ sont donc respectivement mutuellement indépendantes, et par hypothèse au plus une composante est de distribution gaussienne. D'après le théorème 1 on a donc :

$$\mathbf{U}^H \mathbf{U}' = \mathbf{P} \Phi \quad (1.65)$$

où $\mathbf{P} \in \mathbb{R}^{n \times n}$ est une matrice de permutation et $\Phi \in \mathbb{C}^{n \times n}$ est une matrice diagonale à coefficients de module unitaire.

D'après les équations (1.63) et (1.62) on a donc :

$$\mathbf{A}' \mathcal{R}_{\mathbf{s}'\mathbf{s}'}^{-\frac{1}{2}} = \mathbf{W}^\# \mathbf{U} \mathbf{P} \Phi \quad (1.66)$$

$$= \mathbf{A} \mathcal{R}_{\mathbf{s}\mathbf{s}}^{-\frac{1}{2}} \mathbf{P} \Phi \quad (1.67)$$

d'où :

$$\mathbf{A}' = \mathbf{A} \mathcal{R}_{\mathbf{s}\mathbf{s}}^{-\frac{1}{2}} \mathbf{P} \Phi \mathcal{R}_{\mathbf{s}'\mathbf{s}'}^{-\frac{1}{2}} \quad (1.68)$$

et donc $\mathbf{A}' \sim \mathbf{A}$. \square

1.3.5 Discussion

Nous avons démontré l'identifiabilité de quatre modèles de sources. La différence fondamentale entre ces modèles concerne l'existence d'une *diversité temporelle et/ou fréquentielle* des sources ou non. La propriété clé qui justifie l'identifiabilité de ces modèles est la notion d'indépendance mutuelle :

- Si les sources possèdent une *diversité temporelle et/ou fréquentielle* alors leur indépendance mutuelle à l'ordre 2 suffit à garantir leur identifiabilité (modèles $\mathbb{M}_1, \mathbb{M}_2, \mathbb{M}_3$),
- Si les sources ne possèdent pas de *diversité temporelle ou fréquentielle* et qu'elles sont modélisées par des séquences i.i.d alors leur indépendance mutuelle au sens strict est une façon de garantir leur identifiabilité, à condition qu'une source au plus soit de densité gaussienne (modèle \mathbb{M}_4).

Par ailleurs, il est important de dissocier l'identifiabilité d'un modèle des algorithmes construits sur la base des hypothèses définissant ce modèle. Le fait qu'un modèle soit identifiable ne signifie pas qu'*en pratique* on sache effectivement séparer des mélanges de sources répondant aux hypothèses de ce modèle.

Ainsi, la preuve de l'identifiabilité de \mathbb{M}_4 n'est pas constructive et nous verrons dans la partie 2.1 comment le problème de la SAS sous les hypothèses du modèle \mathbb{M}_4 est ramené à un problème d'Analyse en Composantes Indépendantes.

En revanche, nous avons souligné au paragraphe 1.3.1 que la preuve de l'identifiabilité de \mathbb{M}_1 (et donc de \mathbb{M}_2 et \mathbb{M}_3) est constructive à condition de pouvoir estimer \mathbf{W} et $\mathcal{S}_{\mathbf{xx}}[t, f]$ (et en négligeant le bruit), ce qui dans le cadre général de sources non-stationnaires colorées n'est pas trivial. C'est pourquoi l'intérêt des modèles \mathbb{M}_2 et \mathbb{M}_3 provient précisément du fait que sous leurs hypothèses moins générales que celles du modèle \mathbb{M}_1 , l'estimation *en pratique* du SWVS des observations $\mathcal{S}_{\mathbf{xx}}[t, f]$ est beaucoup plus aisée. Ainsi nous présenterons dans le paragraphe 2.2 la méthode SOBI [BAMCM97] construite sur les hypothèses du modèle \mathbb{M}_3 . De la même façon, nous présenterons dans le paragraphe 3.1 une méthode de séparation proposée par Pham & Cardoso [PC01] construite sur les hypothèses du modèle \mathbb{M}_2 . Le paragraphe 3.2 est quant à lui consacré à la construction d'un algorithme basé sur les hypothèses plus générales du modèle \mathbb{M}_1 .

Enfin, nous soulignons dès à présent que le fait de construire un algorithme à partir des hypothèses d'un modèle choisi \mathbb{M} ne signifie pas que l'algorithme est voué à l'échec lorsque les sources qu'il doit estimer ne satisfont plus \mathbb{M} . La robustesse de l'algorithme vis à vis de l'écart des sources au modèle sur lequel l'algorithme a été construit reste alors à étudier. Ceci sera illustré dans les simulations de la partie 3.3.

Chapitre 2

Algorithmes standard de séparation en mélange linéaire instantané

L'objectif commun des différentes méthodes que nous allons présenter est l'estimation d'une matrice équivalente à \mathbf{A} (donc à gains et permutations près sur les colonnes) ou d'une matrice équivalente à $\mathbf{A}^\#$ (cette fois à gains et permutations près sur les lignes). Plusieurs des méthodes exposées par la suite procèdent par estimation d'une *matrice de séparation* $\mathbf{B} \in \mathbb{C}^{n \times m}$ telle que

$$\mathbf{B} = \mathbf{P} \mathbf{D} \mathbf{A}^\# \quad (2.1)$$

où $\mathbf{P} \in \mathbb{R}^{n \times n}$ est une matrice de permutation et $\mathbf{D} \in \mathbb{C}^{n \times n}$ est une matrice diagonale inversible. Notre estimation des sources sera alors :

$$\hat{\mathbf{s}} = \mathbf{B} \mathbf{x} \quad (2.2)$$

$$= \mathbf{P} \mathbf{D} \mathbf{s} + \mathbf{B} \mathbf{n} \quad (2.3)$$

Ainsi, dans le cadre de nos travaux, les estimations des sources seront bruitées. Dans le contexte de la séparation de sources audio, une étape de débruitage supplémentaire de $\hat{\mathbf{s}}$ pourrait être envisagée en utilisant par exemple une méthode de restauration audio [WG03]. Cependant cette perspective n'est pas étudiée ici.

2.1 Séparation de sources i.i.d (modèle \mathbb{M}_4)

Nous proposons dans un premier temps de nous intéresser aux méthodes de séparation de mélanges de sources répondant au modèle \mathbb{M}_4 , dans la mesure où elles constituent l'approche «historique» de la séparation de sources.

Supposons $\mathbf{s} \in \mathbb{M}_4$. Supposons par ailleurs dans un premier temps $\mathbf{n} = \mathbf{0}$. On a alors :

$$\mathbf{x} = \mathbf{A} \mathbf{s} \quad (2.4)$$

D'où, $\forall t \in \llbracket 0, T-1 \rrbracket$:

$$\mathbf{x}[t] = \sum_{j=1}^n \mathbf{a}_j s_j[t] \quad (2.5)$$

où \mathbf{a}_j désigne la j^{eme} colonne de \mathbf{A} . Les sources étant identiquement distribuées dans le temps, les observations sont donc aussi identiquement distribuées dans le temps :

$$\forall t \neq t', \quad p(\mathbf{x}[t]) = p(\mathbf{x}[t']) \quad (2.6)$$

Chaque source s_i étant indépendamment distribuée dans le temps et toutes les sources étant de surcroît mutuellement indépendantes, les observations sont indépendamment distribuées dans le

temps :

$$\forall t \neq t', \quad p(\mathbf{x}[t], \mathbf{x}[t']) = p\left(\sum_{j=1}^n \mathbf{a}_j s_j[t], \sum_{j=1}^n \mathbf{a}_j s_j[t']\right) \quad (2.7)$$

$$= p(\mathbf{x}[t]) p(\mathbf{x}[t']) \quad (2.8)$$

Les observations sont donc aussi des séquences aléatoires i.i.d, mais elles ne sont plus mutuellement indépendantes. L'objectif de la SAS devient alors de trouver une matrice de séparation $\mathbf{B} \in \mathbb{C}^{n \times m}$ telle que les composantes de $\mathbf{y} = \mathbf{B} \mathbf{x}$ soient mutuellement indépendantes. On aura alors $(\mathbf{B}^\#, \mathbf{y}) \in \mathbb{I}_{(\mathbf{x}-\mathbf{n})}$ et $\mathbf{y} \in \mathbb{M}_4$, et donc d'après la propriété 8 :

$$(\mathbf{B}^\#, \mathbf{y}) \sim (\mathbf{A}, \mathbf{s}) \quad (2.9)$$

On s'aperçoit donc que l'objectif de la SAS, dans le cas de sources supposées i.i.d et mutuellement indépendantes, est de recouvrer l'indépendance des composantes de \mathbf{y} et non leur propriété i.i.d qui est de toute façon conservée par la multiplication de \mathbf{s} par les matrices \mathbf{A} et \mathbf{B} .

Dans un premier temps nous montrerons comment écrire un estimateur au sens du Maximum de Vraisemblance (MV) de \mathbf{A} , en utilisant une distribution *hypothétique* des sources. Nous verrons ensuite comment il est possible, partant du MV, d'aboutir à des méthodes de type entropique/information mutuelle s'affranchissant de toute hypothèse sur la distribution des sources.

En pratique, le bruit \mathbf{n} sur les observations est traité comme un «paramètre de nuisance», c-à-d qu'il n'est pas pris en compte dans l'écriture des méthodes : on cherchera la matrice de séparation \mathbf{B} telle que les composantes de \mathbf{y} soient *les plus indépendantes possibles*, au sens d'une certaine *mesure d'indépendance* entre les composantes de \mathbf{y} , et le problème de la SAS devient alors celui de l'Analyse en Composantes Indépendantes (ACI) [Com94]. Dans les paragraphes qui suivent, et jusqu'à indication contraire, on supposera donc $\mathbf{n} = \mathbf{0}$.

Le fil conducteur des parties suivantes est celui adopté par Cardoso dans les articles [Car98a, Car99].

2.1.1 Indépendance et décorrélation

Blanchiment spatial

Étant données des observations \mathbf{x} nous avons donc ramené notre problème de SAS à un problème d'ACI qui consiste à trouver \mathbf{B} telle que les composantes de $\mathbf{y} = \mathbf{B} \mathbf{x}$ soient indépendantes. A cette fin une première idée consiste à réaliser la *décorrélation* des observations, c-à-d, à les rendre *indépendantes à l'ordre 2*. Ceci revient à calculer \mathbf{B} telle que $\forall i \neq j \in \llbracket 1, n \rrbracket^2, \mathbb{E}\{y_i[\cdot] \bar{y}_j[\cdot]\} = 0$, soit encore, en posant $\mathcal{R}_{\mathbf{y}\mathbf{y}} = \mathcal{R}_{\mathbf{y}\mathbf{y}}[t, 0]$:

$$\mathcal{R}_{\mathbf{y}\mathbf{y}} = \mathbf{D} \quad (2.10)$$

où $\mathbf{D} \in \mathbb{R}^{n \times n}$ est une matrice diagonale > 0 . En posant $\mathcal{R}_{\mathbf{x}\mathbf{x}} = \mathcal{R}_{\mathbf{x}\mathbf{x}}[t, 0]$, cela revient à chercher \mathbf{B} telle que

$$\mathbf{B} \mathcal{R}_{\mathbf{x}\mathbf{x}} \mathbf{B}^H = \mathbf{D} \quad (2.11)$$

En tenant compte du caractère hermitien de $\mathcal{R}_{\mathbf{x}\mathbf{x}}$, la décorrélation n'offre donc que $n(n-1)/2$ équations (non-linéaires) pour mn paramètres à estimer. Ceci s'avère être insuffisant pour résoudre notre problème, ce qui signifie que l'indépendance à l'ordre 2 ne suffit pas.

Cependant, nous verrons que la décorrélation préalable des observations peut simplifier l'écriture de plusieurs méthodes d'ACI. Cette opération de décorrélation, assortie de la normalisation des variances des séquences, est couramment appelé le *blanchiment spatial*. Le blanchiment spatial consiste donc à calculer une matrice $\mathbf{W} \in \mathbb{C}^{n \times m}$ telle que $\mathbf{z} \stackrel{\text{def}}{=} \mathbf{W} \mathbf{x}$ ait des composantes de variance unitaire mutuellement décorréliées, c-à-d, en posant $\mathcal{R}_{\mathbf{z}\mathbf{z}} = \mathcal{R}_{\mathbf{z}\mathbf{z}}[t, 0]$:

$$\mathcal{R}_{\mathbf{z}\mathbf{z}} = \mathbf{I}_n \quad (2.12)$$

Avec $\mathbf{z} = \mathbf{W} \mathbf{A} \mathbf{s}$, \mathbf{W} doit donc vérifier :

$$\mathbf{W} \mathbf{A} \mathcal{R}_{\mathbf{ss}} \mathbf{A}^H \mathbf{W}^H = \mathbf{I}_n \quad (2.13)$$

En posant $\mathbf{U} = \mathbf{W} \mathbf{A} \mathcal{R}_{\mathbf{ss}}^{-\frac{1}{2}}$, \mathbf{U} est une matrice orthonormale et on a :

$$\mathbf{z} = \mathbf{W} \mathbf{A} \mathbf{x} \quad (2.14)$$

$$= \mathbf{U} \mathcal{R}_{\mathbf{ss}}^{-\frac{1}{2}} \mathbf{s} \quad (2.15)$$

Ceci signifie qu'après blanchiment, \mathbf{z} est égal aux sources normalisées $\mathcal{R}_{\mathbf{ss}}^{-\frac{1}{2}} \mathbf{s}$ à un «facteur» orthonormal \mathbf{U} près. En outre, d'après la propriété 3 :

$$\mathbf{U}^H \mathbf{W} = \mathcal{R}_{\mathbf{ss}}^{-\frac{1}{2}} \mathbf{A}^\# \quad (2.16)$$

et

$$\mathbf{W}^\# \mathbf{U} = \mathbf{A} \mathcal{R}_{\mathbf{ss}}^{\frac{1}{2}} \quad (2.17)$$

Le blanchiment spatial est donc un cas particulier d'orthonormalisation, qui permet de résoudre le problème de la séparation des mélanges linéaires instantanés en deux étapes : 1) estimation d'une matrice de blanchiment spatial \mathbf{W} , 2) estimation du facteur orthonormal \mathbf{U} manquant.

Les équations (2.16) et (2.17) montrent qu'il n'est possible d'estimer \mathbf{A} (ou $\mathbf{A}^\#$) qu'à la matrice diagonale $\mathcal{R}_{\mathbf{ss}}^{\frac{1}{2}}$ près, qui contient les écarts types inconnus des sources. La matrice $\mathcal{R}_{\mathbf{ss}}^{\frac{1}{2}}$ caractérise donc l'indétermination sur le gain de la SAS. Nous verrons ultérieurement qu'il nous sera possible d'estimer \mathbf{U} uniquement à des permutations de colonnes près, ce qui caractérisera l'indétermination sur l'ordre des sources. Afin de lever l'indétermination sur le gain on supposera jusqu'à indication contraire :

$$\mathcal{R}_{\mathbf{ss}} = \mathbf{I}_n \quad (2.18)$$

Blanchiment spatial en pratique

On a :

$$\mathcal{R}_{\mathbf{xx}} = \mathbf{A} \mathcal{R}_{\mathbf{ss}} \mathbf{A}^H \quad (2.19)$$

Une matrice \mathbf{W} vérifiant l'équation (2.13) peut alors être calculée d'après les éléments propres de $\mathcal{R}_{\mathbf{xx}}$ comme indiqué au paragraphe 1.2. En pratique, en supposant les signaux ergodiques, $\mathcal{R}_{\mathbf{xx}}$ est estimé par :

$$\hat{\mathcal{R}}_{\mathbf{xx}} = \frac{1}{T} \sum_{t=0}^{T-1} \mathbf{x}[t] \mathbf{x}[t]^H \quad (2.20)$$

Il est à noter que l'erreur sur l'estimation de \mathbf{W} ne peut pas être compensée par l'estimation ultérieure de \mathbf{U} . L'indice de performance des algorithmes de séparation ayant recours au blanchiment spatial préalable des observations possède ainsi une borne inférieure quantifiable [Car94]. Le blanchiment spatial demeure toutefois une opération utile qui ramène le problème de la SAS à la recherche d'une matrice orthonormale, ce dont on tirera parti dans l'élaboration des méthodes de séparation présentées par la suite.

2.1.2 Quelques méthodes d'ACI en théorie

Nous allons présenter dans ce paragraphe plusieurs méthodes d'ACI en théorie. Afin de simplifier la présentation des méthodes nous nous placerons dans le cas $m = n$ et dans le cas où les sources et la matrice de mélange sont à valeurs réelles. La plupart des méthodes présentées peuvent être généralisées au cas sur-déterminé ($m > n$) et au cas complexe. Par exemple, certaines méthodes requièrent un blanchiment spatial des observations afin de se ramener à l'estimation d'une matrice orthonormale de dimensions $n \times n$. Les dimensions initiales de \mathbf{A} n'ont alors pas d'importance. Le lecteur est invité à se reporter aux articles de référence des méthodes exposées pour en connaître l'exacte portée.

Fonctions de contraste

Aux méthodes d'ACI que nous allons présenter est associée une *fonction de contraste* [Com94]. Une fonction de contraste $\mathcal{C}\{\mathbf{v}\}$ est une fonction de la *densité de probabilité* d'une séquence aléatoire $\mathbf{v} \in \mathbb{R}^{n \times T}$ qui vérifie :

$$\begin{cases} \forall \mathbf{M} \in \mathbb{R}^{n \times n}, & \mathcal{C}\{\mathbf{M}\mathbf{s}\} \geq \mathcal{C}\{\mathbf{s}\} \\ \mathcal{C}\{\mathbf{M}\mathbf{s}\} = \mathcal{C}\{\mathbf{s}\} & \text{ssi } \mathbf{M} \sim \mathbf{I}_n \end{cases} \quad (2.21)$$

Le fait que $\mathcal{C}\{\mathbf{v}\}$ est une fonction de la densité de probabilité de \mathbf{v} (et non d'une réalisation de \mathbf{v}) est caractérisé par l'utilisation de crochets $\{\cdot\}$ (au lieu de parenthèses (\cdot)).

Ainsi, pour une fonction de contraste $\mathcal{C}\{\mathbf{v}\}$ donnée, les méthodes que nous allons présenter consisteront à minimiser suivant \mathbf{B} la fonction $\mathcal{C}\{\mathbf{B}\mathbf{x}\} = \mathcal{C}\{\mathbf{B}\mathbf{A}\mathbf{s}\}$. Si une matrice \mathbf{B} minimise $\mathcal{C}\{\mathbf{B}\mathbf{x}\}$ on est alors assuré d'avoir $\mathbf{B}\mathbf{A} \sim \mathbf{I}_n$, ce qui signifie dans le cas $m = n$ que \mathbf{B} est égale à \mathbf{A}^{-1} à gains et permutations près sur les lignes, et qui s'avère être notre objectif.

Par la suite on notera $\mathbf{y} = \mathbf{B}\mathbf{x}$ et on pourra parler de la minimisation de $\mathcal{C}\{\mathbf{y}\}$ par rapport à \mathbf{B} .

Le fait qu'une méthode de séparation soit basée sur la minimisation d'une fonction de contraste $\mathcal{C}\{\mathbf{y}\} = \mathcal{C}\{\mathbf{B}\mathbf{x}\}$ lui confère la propriété d'*équivalence* [CL96]. Cette propriété signifie que les performances de la méthode ne dépendent pas de la matrice de mélange \mathbf{A} .

Il est important de rappeler que, dans le cas de sources supposées i.i.d et mutuellement indépendantes, les séquences \mathbf{s} , \mathbf{x} et \mathbf{y} sont des séquences i.i.d, ce qui signifie que leurs densités de probabilités $p_{\mathbf{s}}$, $p_{\mathbf{x}}$ et $p_{\mathbf{y}}$ ne dépendent respectivement que des densités de probabilité $p_{\mathbf{s}[\cdot]}$, $p_{\mathbf{x}[\cdot]}$ et $p_{\mathbf{y}[\cdot]}$ de n'importe quel échantillon $\mathbf{s}[\cdot]$, $\mathbf{x}[\cdot]$ et $\mathbf{y}[\cdot]$. Ceci signifie en particulier que, dans notre contexte, une fonction de contraste $\mathcal{C}\{\mathbf{y}\}$ s'avère être plus simplement une fonction de $p_{\mathbf{y}[\cdot]}$.

Fonctions de contrastes orthonormales

Nous verrons que la minimisation d'une fonction de contraste revient à rendre les composantes de $\mathbf{y} = \mathbf{B}\mathbf{x}$ les plus indépendantes possibles au sens de la *mesure d'indépendance* définie par la fonction de contraste. Dans la mesure où en pratique l'indépendance «exacte» ne peut être recouvrée, si la décorrélation mutuelle des composantes de \mathbf{y} est souhaitée alors il faut explicitement la prendre en compte dans la minimisation de la fonction de contraste. On parlera alors de minimisation de $\mathcal{C}\{\mathbf{y}\}$ *sous contrainte de blancheur* ou encore de minimisation du *contraste orthonormal* $\mathcal{C}^\circ\{\mathbf{y}\}$. Cela pourra correspondre à deux options :

- ou bien un blanchiment spatial préalable des observations aura été réalisé de sorte que l'estimation de la matrice de séparation \mathbf{B} ait été ramenée à l'estimation d'une matrice de séparation \mathbf{V} orthonormale telle que $\mathcal{C}\{\mathbf{y}\} = \mathcal{C}\{\mathbf{V}\mathbf{z}\}$,

- ou bien la minimisation du contraste $\mathcal{C}\{\mathbf{y}\} = \mathcal{C}\{\mathbf{B}\mathbf{x}\}$ est réalisée sous la contrainte $\mathcal{R}_{\mathbf{y}\mathbf{y}} = \mathbf{I}_n$.

Puisque les fonctions de contraste dépendent de la densité de probabilité d'une séquence aléatoire, il apparaît que leur estimation d'après des données empiriques n'est pas triviale. Ce problème sera l'objet du paragraphe 2.1.3, où l'implémentation pratique des différentes méthodes d'ACI présentées en théorie sera considérée. Nous étudierons en particulier l'algorithme JADE.

Approche du Maximum de Vraisemblance

Nous présentons dans ce paragraphe une solution au sens du Maximum de Vraisemblance (MV) du problème de la séparation de sources [PG97, Car97, Car98a]. Soit $q(\mathbf{s}[\cdot])$ une distribution *hypothétique* d'un échantillon quelconque $\mathbf{s}[\cdot]$ des sources.

Les sources étant mutuellement indépendantes on écrira :

$$q(\mathbf{s}[\cdot]) = \prod_{i=1}^n q_i(s_i[\cdot]) \quad (2.22)$$

Avec $\mathbf{x}[\cdot] = \mathbf{A}\mathbf{s}[\cdot]$, la densité de probabilité d'un échantillon quelconque des observations s'écrit [Pap84] :

$$p_{\mathbf{x}[\cdot]}(\mathbf{x}[\cdot]; \mathbf{A}, q) = |\det(\mathbf{A})|^{-1} q(\mathbf{A}^{-1}\mathbf{x}[\cdot]) \quad (2.23)$$

Les observations \mathbf{x} étant indépendamment distribuées, la densité de probabilité de \mathbf{x} s'écrit :

$$p_{\mathbf{x}}(\mathbf{x}; \mathbf{A}, q) = \prod_{t=0}^{T-1} |\det(\mathbf{A})|^{-1} q(\mathbf{A}^{-1}\mathbf{x}[t]) \quad (2.24)$$

En notant $\mathbf{B} = \mathbf{A}^{-1}$, la log-vraisemblance de \mathbf{B} (et q) par rapport à \mathbf{x} s'écrit donc :

$$L(\mathbf{B}, q) \stackrel{\text{def}}{=} \frac{1}{T} \log p_{\mathbf{x}}(\mathbf{x}; \mathbf{B}, q) \quad (2.25)$$

$$= \log |\det(\mathbf{B})| + \frac{1}{T} \sum_{t=0}^{T-1} \log q(\mathbf{B}\mathbf{x}[t]) \quad (2.26)$$

Asymptotiquement, lorsque $T \rightarrow +\infty$, le critère du MV peut être formulé comme une divergence de Kullback entre deux densités de probabilité. Soient \mathbf{v} et \mathbf{w} deux vecteurs aléatoires sur $\mathbb{R}^{n \times 1}$. La divergence de Kullback de \mathbf{v} à \mathbf{w} est définie par :

$$\mathcal{K}\{\mathbf{v} \parallel \mathbf{w}\} \stackrel{\text{def}}{=} \int_{\mathbf{u}} p_{\mathbf{v}}(\mathbf{u}) \ln \frac{p_{\mathbf{v}}(\mathbf{u})}{p_{\mathbf{w}}(\mathbf{u})} d\mathbf{u} \quad (2.27)$$

Une propriété importante de la divergence de Kullback est qu'elle est positive, nulle ssi \mathbf{v} et \mathbf{w} ont la même distribution. Soit \mathbf{s}' une séquence aléatoire i.i.d de densité $p_{\mathbf{s}'[\cdot]} = q$. Alors on peut montrer [Car97] :

$$L(\mathbf{B}, q) \xrightarrow{T \rightarrow +\infty} -\mathcal{K}\{\mathbf{B}\mathbf{x}[\cdot] \parallel \mathbf{s}'[\cdot]\} + \text{constante} \quad (2.28)$$

Pour une distribution q fixée, on peut ainsi associer au MV le contraste suivant :

$$\mathcal{C}_{MV}\{\mathbf{y}\} = \mathcal{K}\{\mathbf{B}\mathbf{x}[\cdot] \parallel \mathbf{s}'[\cdot]\} \quad (2.29)$$

Sous cet aspect, l'estimateur du MV offre une nouvelle interprétation : il consiste à trouver la matrice de séparation \mathbf{B} qui minimise la divergence de Kullback entre la densité de $\mathbf{y}[\cdot] = \mathbf{B}\mathbf{x}[\cdot]$ et la densité hypothétique q des sources. En d'autres termes on cherche \mathbf{B} qui fournit des estimations des sources de distribution se rapprochant le plus, au sens de la divergence de Kullback,

des hypothèses qui ont été faites sur la densité des sources.

Pour une distribution q fixée, ayant levé l'indétermination sur le gain de la SAS en fixant $\mathcal{R}_{ss} = \mathbf{I}_n$, $L(\mathbf{B}, q)$ possède plusieurs maxima correspondant aux permutations de lignes possibles de \mathbf{B} . Pham & Garat montrent dans [PG97] que le gradient de $L(\mathbf{B}, q)$ par rapport à \mathbf{B} peut être calculé explicitement et qu'il dépend des *fonctions score* des sources. $\forall i \in \llbracket 1, n \rrbracket$, la fonction score de la source s_i , de densité de probabilité q_i , est définie par :

$$\forall u \in \mathbb{R}, \quad \varphi_i(u) \stackrel{\text{def}}{=} \frac{\delta \ln q_i}{\delta u}(u) \quad (2.30)$$

$$= \frac{1}{q_i(u)} \frac{\delta q_i}{\delta u}(u) \quad (2.31)$$

On définit le *vecteur score* φ de \mathbf{s} par :

$$\forall \mathbf{u} = [u_1, \dots, u_n]^T \in \mathbb{R}^{n \times 1}, \quad \varphi(\mathbf{u}) \stackrel{\text{def}}{=} [\varphi_1(u_1), \dots, \varphi_n(u_n)]^T \quad (2.32)$$

Le gradient de $L(\mathbf{B}, q)$ par rapport à \mathbf{B} s'écrit alors [PG97] :

$$\frac{\delta L}{\delta \mathbf{B}}(\mathbf{B}, q) = \left(\frac{1}{T} \sum_{t=0}^{T-1} H(\mathbf{B} \mathbf{x}[t]) \right) \mathbf{B}^{-T} \quad (2.33)$$

où $H(\mathbf{u})$ est la fonction de $\mathbb{R}^{n \times 1} \rightarrow \mathbb{R}^{n \times n}$ définie par :

$$H(\mathbf{u}) \stackrel{\text{def}}{=} \varphi(\mathbf{u}) \mathbf{u}^T - \mathbf{I}_n \quad (2.34)$$

Si la distribution des sources est connue, le critère $L(\mathbf{B}, q)$ peut être facilement maximisé selon \mathbf{B} en utilisant une méthode de type de descente de gradient, en-ligne ou hors-ligne, en utilisant l'expression (2.33) du gradient. Cependant, dans cette formulation, le calcul du gradient à chaque itération nécessite l'inversion de la matrice \mathbf{B} , ce qui s'avère être une opération coûteuse et qui prive la méthode de la propriété d'équivariance. C'est pourquoi, il est préconisé d'utiliser l'expression du critère de MV sous forme de contraste et d'utiliser le gradient de $\mathcal{C}_{MV}\{\mathbf{y}\}$ «relatif» à la structure particulière de l'espace défini par $\mathbf{y} = \mathbf{B} \mathbf{x}$. Le concept de *gradient relatif* $\nabla_{\mathbf{B}} \mathcal{C}_{MV}\{\mathbf{y}\}$ de la fonction de contraste $\mathcal{C}\{\mathbf{y}\}$ sous la structure $\mathbf{y} = \mathbf{B} \mathbf{x}$ s'écrit :

$$\nabla_{\mathbf{B}} \mathcal{C}_{MV}\{\mathbf{B} \mathbf{x}\} = \frac{\delta \mathcal{C}\{\mathbf{B} \mathbf{x}\}}{\delta \mathbf{B}} \mathbf{B}^T \quad (2.35)$$

En utilisant par exemple l'expression asymptotique de l'équation (2.33), on peut montrer que le gradient relatif de la fonction de contraste associée au critère de MV s'écrit :

$$\nabla_{\mathbf{B}} \mathcal{C}_{MV}\{\mathbf{y}\} = \mathbb{E}\{H(\mathbf{y}[\cdot])\} \quad (2.36)$$

En pratique, la minimisation de $\mathcal{C}_{MV}\{\mathbf{y}\}$ peut alors être réalisée par descente de gradient relatif en remplaçant l'espérance dans l'expression (2.36) par une moyenne temporelle. La trame de la méthode hors-ligne peut alors s'écrire :

$$\text{Initialisation : } \mathbf{B} = \mathbf{I}_n, \quad \mathbf{y} = \mathbf{x} \quad (2.37)$$

$$\hat{H} \leftarrow \frac{1}{T} \sum_{t=0}^{T-1} H(\mathbf{y}[t]) \quad (2.38)$$

$$\mathbf{B} \leftarrow (\mathbf{I}_n - \mu \hat{H}) \mathbf{B} \quad (2.39)$$

$$\mathbf{y} \leftarrow \mathbf{B} \mathbf{y} \quad (2.40)$$

où μ désigne le pas de descente, éventuellement variable. La mise à jour peut aussi être réalisée directement sur les données sans l'intermédiaire de la matrice \mathbf{B} :

$$\text{Initialisation :} \quad \mathbf{y} = \mathbf{x} \quad (2.41)$$

$$\hat{H} \leftarrow \frac{1}{T} \sum_{t=0}^{T-1} H(\mathbf{y}[t]) \quad (2.42)$$

$$\mathbf{y} \leftarrow (\mathbf{I}_n - \mu \hat{H}) \mathbf{y} \quad (2.43)$$

Une version en-ligne est possible en utilisant une méthode de descente de *gradient stochastique*. La méthode de descente de gradient stochastique consiste à «omettre» l'opérateur d'espérance dans la formulation du gradient relatif (2.36) et à utiliser simplement $H(\mathbf{B} \mathbf{x}[t])$ dans la mise à jour de \mathbf{B} à l'arrivée d'un échantillon $\mathbf{x}[t]$ à l'instant t . La méthode possède alors la formulation en-ligne suivante :

$$\text{Initialisation :} \quad \mathbf{B}[0] = \mathbf{I}_n \quad (2.44)$$

$$\mathbf{y}[t] \leftarrow \mathbf{B}[t] \mathbf{x}[t] \quad (2.45)$$

$$\mathbf{B}[t+1] \leftarrow (\mathbf{I}_n - \mu[t] H(\mathbf{y}[t])) \mathbf{B}[t] \quad (2.46)$$

La méthode du MV que nous venons de décrire repose sur l'utilisation d'une densité hypothétique q des sources, ou plus précisément d'un vecteur score hypothétique φ des sources. De nombreux travaux ont montré qu'il n'est pas nécessaire de connaître de manière précise les fonctions score réelles des sources pour que les maxima du critère de MV correspondent à la matrice de mélange réelle à permutations près. Une illustration simple de ce phénomène est présentée dans [Car98a]. La «marge d'erreur» autorisée sur la spécification de φ peut être quantifiée, ceci est l'objet par exemple des travaux [CL96, ACC97].

On remarquera que la condition d'annulation du gradient relatif de $\mathcal{C}_{MV}\{\mathbf{y}\}$ s'écrit :

$$\mathbb{E}\{\varphi(\mathbf{y}[\cdot]) \mathbf{y}[\cdot]^T\} = \mathbf{I}_n \quad (2.47)$$

Les fonctions score étant non-linéaires (sauf pour des variables aléatoires gaussiennes...) on voit que le MV revient à réaliser une décorrélation «non-linéaire» des observations. Cette idée est la clé des travaux «pionniers» de Jutten & Héroult en SAS [JH91].

Enfin, la minimisation de $\mathcal{C}_{MV}\{\mathbf{y}\}$ sous contrainte de blancheur est possible sans blanchiment spatial préalable en remplaçant dans les algorithmes de minimisation H par H° définie par [CL96] :

$$H^\circ(\mathbf{u}) = \mathbf{u} \mathbf{u}^T - \mathbf{I}_n + \varphi(\mathbf{u}) \mathbf{u}^T - \mathbf{u} \varphi(\mathbf{u})^T \quad (2.48)$$

Maximum de vraisemblance et Infomax

Le contraste $\mathcal{C}_{MV}\{\mathbf{y}\}$ a été proposé par ailleurs par Bell et Sejnowski via une approche réseau de neurones [BS95], donnant à leur méthode le nom d'*Infomax*. Soient g_1, \dots, g_n les fonctions de répartitions de q_1, \dots, q_n définies par :

$$g_i(u) = \int_{-\infty}^u q_i(v) dv \quad \in [0, 1] \quad (2.49)$$

Pour $\mathbf{u} = [u_1, \dots, u_n]^T \in \mathbb{R}^{n \times 1}$ notons $g(\mathbf{u}) = [g_1(u_1), \dots, g_n(u_n)]$. La fonction de contraste de la méthode Infomax est définie par :

$$\mathcal{C}_{Info}\{\mathbf{y}\} \stackrel{\text{def}}{=} -\mathcal{H}\{g(\mathbf{y}[\cdot])\} \quad (2.50)$$

où $\mathcal{H}\{\mathbf{v}\}$ désigne l'entropie différentielle du vecteur aléatoire \mathbf{v} sur $\mathbb{R}^{n \times 1}$, définie par :

$$\mathcal{H}\{\mathbf{v}\} \stackrel{\text{def}}{=} - \int_{\mathbf{u}} p_{\mathbf{v}}(\mathbf{u}) \ln p_{\mathbf{v}}(\mathbf{u}) d\mathbf{u} \quad (2.51)$$

L'idée sous-jacente au contraste $\mathcal{C}_{\text{Info}}\{\mathbf{y}\}$ est la suivante : si comme précédemment le vecteur aléatoire $\mathbf{s}'[\cdot]$ a pour densité q , alors, puisque $\forall i \in \llbracket 1, n \rrbracket$, g_i est la fonction de répartition de q_i , le vecteur aléatoire $g(\mathbf{s}'[\cdot])$ est uniformément distribué sur $[0, 1]^n$. Or, la densité uniforme sur $[0, 1]^n$ est la densité possédant la plus haute entropie de toutes les densités définies sur $[0, 1]^n$. Donc $g(\mathbf{B}\mathbf{x}[\cdot])$ a la plus forte entropie si $\mathbf{B}\mathbf{x}[\cdot] \sim \mathbf{s}'[\cdot]$, d'où le contraste $\mathcal{C}_{\text{Info}}\{\mathbf{y}\}$. La méthode Infomax consiste donc à chercher \mathbf{B} telle que la densité de $\mathbf{B}\mathbf{x}[\cdot]$ se rapproche de la densité hypothétique q au sens d'une mesure entropique. L'équivalence entre $\mathcal{C}_{\text{Info}}\{\mathbf{y}\}$ et $\mathcal{C}_{\text{MV}}\{\mathbf{y}\}$ est démontrée dans [Car97].

Vers des approches sans modèle des sources

L'avantage principal de la méthode du MV est qu'elle est simple à mettre en oeuvre ; le gradient relatif de la fonction de contraste $\mathcal{C}_{\text{MV}}\{\mathbf{y}\}$ se calcule explicitement ce qui permet de minimiser le critère par une méthode de gradient simple.

L'inconvénient majeur est qu'un *a priori* sur la densité des sources est nécessaire. Nous allons voir dans ce paragraphe comment il est possible de s'affranchir de tout *a priori* sur les sources, au détriment d'une mise en oeuvre en pratique plus délicate.

La fonction de contraste du MV dépend explicitement de \mathbf{B} et implicitement de la distribution hypothétique q choisie. Nous allons voir qu'il est possible de minimiser $\mathcal{C}_{\text{MV}}\{\mathbf{y}\}$ suivant \mathbf{B} et q .

Soit, à \mathbf{B} fixé, $\tilde{\mathbf{y}}[\cdot]$ le vecteur aléatoire dont les composantes $\tilde{y}_1[\cdot], \dots, \tilde{y}_n[\cdot]$ sont mutuellement indépendantes et telles que la densité de chaque composante $\tilde{y}_i[\cdot]$ est égale à la densité de la $i^{\text{ème}}$ marginale $y_i[\cdot]$ de $\mathbf{y}[\cdot]$. Alors, les composantes $\mathbf{s}'[\cdot]$ étant indépendantes, on a [CT91] :

$$\mathcal{K}\{\mathbf{y}[\cdot] \parallel \mathbf{s}'[\cdot]\} = \mathcal{K}\{\mathbf{y}[\cdot] \parallel \tilde{\mathbf{y}}[\cdot]\} + \mathcal{K}\{\tilde{\mathbf{y}}[\cdot] \parallel \mathbf{s}'[\cdot]\} \quad (2.52)$$

$\mathcal{K}\{\mathbf{y}[\cdot] \parallel \tilde{\mathbf{y}}[\cdot]\}$ ne dépendant pas de \mathbf{s}' (c-à-d, de q), la minimisation de $\mathcal{K}\{\mathbf{y}[\cdot] \parallel \mathbf{s}'[\cdot]\}$ suivant q est équivalente à la minimisation de $\mathcal{K}\{\tilde{\mathbf{y}}[\cdot] \parallel \mathbf{s}'[\cdot]\}$ suivant q . À \mathbf{B} fixé, $\mathcal{K}\{\tilde{\mathbf{y}}[\cdot] \parallel \mathbf{s}'[\cdot]\}$ est minimal (égal à zéro) en choisissant $\mathbf{s}'[\cdot] = \tilde{\mathbf{y}}[\cdot]$. Ainsi, la maximisation de la vraisemblance $L(\mathbf{B}, q)$ par rapport à \mathbf{B} et q revient à la minimisation du contraste :

$$\mathcal{C}_{\text{IM}}\{\mathbf{y}\} \stackrel{\text{def}}{=} \mathcal{K}\{\mathbf{y}[\cdot] \parallel \tilde{\mathbf{y}}[\cdot]\} \quad (2.53)$$

$\mathcal{C}_{\text{IM}}\{\mathbf{y}\}$ n'est rien d'autre que l'information mutuelle du vecteur aléatoire \mathbf{y} , définie comme la divergence de Kullback entre la densité d'un vecteur aléatoire et la densité du même vecteur mais avec des composantes mutuellement indépendantes. L'information mutuelle est donc une mesure directe de l'indépendance des composantes d'un vecteur aléatoire.

Sous contrainte de blancheur, le contraste $\mathcal{C}_{\text{IM}}\{\mathbf{y}\}$ peut prendre une forme plus simple. En effet, on peut montrer que sous la contrainte $\mathbb{E}\{\mathbf{y}[\cdot] \mathbf{y}[\cdot]^T\} = \mathbf{I}_n$, l'information mutuelle est égale, à une constante près à la somme des entropies des marginales de \mathbf{y} [Com94]. D'où le contraste :

$$\mathcal{C}_{\text{IM}}^{\circ}\{\mathbf{y}\} \stackrel{\text{c}}{=} \sum_{i=1}^n \mathcal{H}\{y_i\} \quad (2.54)$$

où la notation $\stackrel{\text{c}}{=}$ signifie l'égalité à une constante près. Le contraste $\mathcal{C}_{\text{IM}}^{\circ}\{\mathbf{y}\}$ offre une nouvelle interprétation de l'ACI : le fait de mélanger des sources tend à «gaussianiser» la densité de chaque observation et donc, dans la mesure où la densité gaussienne possède la plus forte entropie des densités de même variance, à augmenter leur entropie. Il est alors légitime de rechercher

\mathbf{B} comme la matrice de séparation qui décroît au maximum l'entropie des composantes de $\mathbf{y}[\cdot]$, c-à-d qui les écarte le plus de la densité gaussienne. Ceci explique d'une certaine façon pourquoi on ne peut tolérer dans le modèle \mathbb{M}_4 au plus une source gaussienne : dans la mesure où la densité d'une somme de variables aléatoires gaussiennes est elle-même gaussienne, on ne peut pas rendre «moins gaussiennes» des combinaisons linéaires de sources gaussiennes. Ces idées sont largement détaillées dans [HKO01].

Les contrastes $\mathcal{C}_{IM}\{\mathbf{y}[\cdot]\}$ ou $\mathcal{C}_{IM}^\circ\{\mathbf{y}[\cdot]\}$ sont des contrastes «canoniques» dans la mesure où ils expriment explicitement l'objectif de l'ACI qui consiste à recouvrer l'indépendance des composantes de $\mathbf{y}[\cdot]$. Ils ne nécessitent donc pas d'information sur la structure de chacune des sources ; ils n'utilisent que l'information «conjointe» que l'on possède des sources, à savoir leur indépendance mutuelle. Cependant, leur inconvénient majeur est que leur gradient est difficile à exprimer et notamment difficile à estimer en pratique [Pha96]. En particulier, leur gradient ne se résume pas à l'espérance d'une fonction des observations comme pour l'estimateur du MV.

Cependant, nous allons voir qu'une approximation simple à base de cumulants des divers contrastes présentés permet de construire des algorithmes efficaces, et qu'aux méthodes de minimisation par descente de gradient peuvent se substituer des méthodes de minimisation par diagonalisation simultanée. L'avantage des cumulants est que leur estimation en pratique est très simple.

2.1.3 Quelques méthodes d'ACI en pratique

Approximation de la divergence de Kullback par des cumulants

Soient a, b, c et d quatre variables aléatoires *centrées* réelles. Le cumulant d'ordre 4 de ces variables aléatoires est défini par :

$$c_{abcd} \stackrel{\text{def}}{=} E\{abcd\} - E\{ab\}E\{cd\} - E\{ac\}E\{bd\} - E\{ad\}E\{bc\} \quad (2.55)$$

Le cumulant d'ordre 2 de deux variables aléatoires désigne simplement la corrélation de ces variables.

Si les variables aléatoires a, b, c et d peuvent être séparées en deux ensembles de variables aléatoires mutuellement indépendantes alors leur cumulant d'ordre 4 s'annule. Par exemple si a est indépendant de b, c et d alors $c_{abcd} = 0$. Ceci montre que les cumulants d'ordre 4 constituent une façon de tester l'indépendance entre plusieurs variables aléatoires, au delà des cumulants d'ordre 2.

Le cumulant d'ordre 4 c_{aaaa} d'une variable aléatoire a centrée est appelé *kurtosis* et noté :

$$k_a \stackrel{\text{def}}{=} c_{aaaa} \quad (2.56)$$

Si a, b, c et d sont de densité conjointe gaussienne alors leur cumulant d'ordre 4 est nul (plus généralement, tous leurs cumulants d'ordre supérieur à 2 sont nuls).

Si a et b sont deux variables aléatoires de densités «proches» de la gaussienne, on peut écrire à l'aide d'un développement d'Edgeworth [Car99] :

$$\mathcal{K}\{a \parallel b\} \approx \frac{1}{4}(\sigma_a^2 - \sigma_b^2)^2 + \frac{1}{48}(k_a - k_b)^2 \quad (2.57)$$

où l'on a noté $\sigma_a^2 = r_{aa}$ et $\sigma_b^2 = r_{bb}$ les variances de a et b .

Cette approximation est aussi valable pour des vecteurs aléatoires $\mathbf{v} = [v_1, \dots, v_n]^T$ et $\mathbf{w} = [w_1, \dots, w_n]^T$ sur $\mathbb{R}^{n \times 1}$:

$$\mathcal{K}\{\mathbf{v} \parallel \mathbf{w}\} \approx \frac{1}{4} \sum_{ij} (r_{v_i v_j} - r_{w_i w_j})^2 + \frac{1}{48} \sum_{ijkl} (c_{v_i v_j v_k v_l} - c_{w_i w_j w_k w_l})^2 \quad (2.58)$$

Disposant de cette approximation, il est possible d'approximer la fonction de contraste du MV définie par l'équation (2.29). La séquence aléatoire \mathbf{y} étant i.i.d, on utilisera la notation suivante :

$$r_{y_i y_j} = r_{y_i[\cdot] y_j[\cdot]} \quad (2.59)$$

$$c_{y_i y_j y_k y_l} = c_{y_i[\cdot] y_j[\cdot] y_k[\cdot] y_l[\cdot]} \quad (2.60)$$

Cette notation est aussi valable pour les cumulants de $\mathbf{x}[\cdot]$, $\mathbf{s}[\cdot]$ et $\mathbf{s}'[\cdot]$. En notant $r_{s'_i s'_j} = \sigma_{s'_i}^2 \delta_{ij}$ et $c_{s'_i s'_j s'_k s'_l} = k_{s'_i} \delta_{ijkl}$, il vient¹ :

$$\mathcal{C}_{MV}\{\mathbf{y}\} \approx \mathcal{C}_{24}\{\mathbf{y}\} \quad (2.61)$$

avec :

$$\mathcal{C}_{24}\{\mathbf{y}\} \stackrel{\text{def}}{=} \frac{1}{4} \sum_{ij} (r_{y_i y_j} - \sigma_{s'_i}^2 \delta_{ij})^2 + \frac{1}{48} \sum_{ijkl} (c_{y_i y_j y_k y_l} - k_{s'_i} \delta_{ijkl})^2 \quad (2.62)$$

Sous cette approximation, le critère de MV apparaît alors simplement comme une mesure de l'écart entre les cumulants d'ordre de 2 et 4 des composantes de $\mathbf{y}[\cdot]$ et des cumulants hypothétiques de \mathbf{s}' . Minimiser $\mathcal{C}_{24}\{\mathbf{y}\}$ sous contrainte de blancheur revient à mettre un poids infini au premier terme de $\mathcal{C}_{24}\{\mathbf{y}\}$. On a donc :

$$\mathcal{C}_{MV}^\circ\{\mathbf{y}\} \approx \mathcal{C}_4^\circ\{\mathbf{y}\} \quad (2.63)$$

avec :

$$\mathcal{C}_4^\circ\{\mathbf{y}\} \stackrel{\text{def}}{=} \frac{1}{48} \sum_{ijkl} (c_{y_i y_j y_k y_l} - k_{s'_i} \delta_{ijkl})^2 \quad (2.64)$$

Comme précédemment, on peut envisager de minimiser $\mathcal{C}_4^\circ\{\mathbf{y}\}$ par rapport à \mathbf{B} mais aussi par rapport aux kurtosis hypothétiques des sources $\{k_{s'_i}, i \in \llbracket 1, n \rrbracket\}$. À \mathbf{B} fixé, la minimisation de $\mathcal{C}_4^\circ\{\mathbf{y}\}$ par rapport à $\{k_{s'_i}, i \in \llbracket 1, n \rrbracket\}$ donne $k_{s'_i} = k_{y_i}$. Cela nous permet d'aboutir au contraste $\mathcal{C}_{kurt}^\circ\{\mathbf{y}\}$ qui peut être considéré comme une approximation de $\mathcal{C}_{IM}^\circ\{\mathbf{y}\}$ [Com94] :

$$\mathcal{C}_{kurt}^\circ\{\mathbf{y}\} \stackrel{\text{def}}{=} \frac{1}{48} \sum_{ijkl \neq iiii} c_{y_i y_j y_k y_l}^2 \quad (2.65)$$

On peut montrer que si $E\{\mathbf{y}[\cdot] \mathbf{y}[\cdot]^T\} = \mathbf{I}_n$ alors la somme des cumulants au carré $\sum_{ijkl} c_{y_i y_j y_k y_l}^2$ est constante [Com94]. $\mathcal{C}_{kurt}^\circ\{\mathbf{y}\}$ peut donc encore s'écrire :

$$\mathcal{C}_{kurt}^\circ\{\mathbf{y}\} \stackrel{c}{=} -\frac{1}{48} \sum_{i=1}^n k_{y_i}^2 \quad (2.66)$$

Le kurtosis d'une variable aléatoire gaussienne étant nul, le fait de maximiser la somme des kurtosis au carré de $\mathbf{y}[\cdot]$ revient encore une fois à écarter le plus possible les densités des composantes de $\mathbf{y}[\cdot]$ d'une densité gaussienne. Le fait que la somme $\sum_{ijkl} c_{y_i y_j y_k y_l}^2$ soit constante sous contrainte de blancheur signifie que la maximisation de la somme des kurtosis au carré revient en outre à annuler les cumulants croisés de $\mathbf{y}[\cdot]$. La méthode JADE que nous allons maintenant présenter propose de ne minimiser que certaines tranches de cumulants, ceci permettant d'aboutir à une procédure efficace de minimisation par diagonalisation simultanée.

¹Si on impose aux estimations des sources d'être de variance unitaire alors $\forall i, \sigma_{s'_i} = 1$.

L'algorithme JADE

L'algorithme JADE, conçu par Cardoso & Souloumiac [CS93], consiste à minimiser la fonction de contraste suivante :

$$\mathcal{C}_{Jade}^o\{\mathbf{y}\} \stackrel{\text{def}}{=} \sum_{ijk} c_{y_i y_j y_k}^2 \quad (2.67)$$

JADE ne teste donc l'indépendance que sur les tranches de la forme (y_i, y_i, y_j, y_k) . Nous allons ramener la minimisation de cette fonction de contraste à la diagonalisation simultanée d'un ensemble de «matrices de cumulants». L'algorithme JADE est plus généralement applicable à des signaux à valeurs complexes mais nous n'exposons ici que le cas réel.

Soit $\mathbf{v} = [v_1, \dots, v_n]^T$ un vecteur aléatoire sur $\mathbb{R}^{n \times 1}$ et \mathbf{M} une matrice $\in \mathbb{R}^{n \times n}$. On définit la «matrice de cumulants» $\mathcal{Q}_{\mathbf{v}}(\mathbf{M}) \in \mathbb{R}^{n \times n}$ par :

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, \quad [\mathcal{Q}_{\mathbf{v}}(\mathbf{M})]_{ij} = \sum_{kl} c_{v_i v_j v_k v_l} m_{kl} \quad (2.68)$$

Les matrices de cumulants ainsi définies vont nous permettre de ramener quelques tranches de la structure tensorielle des cumulants d'ordre 4 à une structure matricielle plus simple.

Dans le cas réel, les matrices de cumulants peuvent aussi s'écrire explicitement de la façon suivante :

$$\mathcal{Q}_{\mathbf{v}}(\mathbf{M}) = \mathbb{E}\{(\mathbf{v}^T \mathbf{M} \mathbf{v}) \mathbf{v} \mathbf{v}^T\} - \mathcal{R}_{\mathbf{v}\mathbf{v}} \text{trace}(\mathbf{M} \mathcal{R}_{\mathbf{v}\mathbf{v}}) - \mathcal{R}_{\mathbf{v}\mathbf{v}} (\mathbf{M} + \mathbf{M}^T) \mathcal{R}_{\mathbf{v}\mathbf{v}} \quad (2.69)$$

L'équation (2.68) montre que chaque coefficient de la matrice $\mathcal{Q}_{\mathbf{v}}(\mathbf{M})$ est une combinaison linéaire particulière de plusieurs tranches de cumulants. Cependant, la formulation de l'équation (2.69) montre qu'il n'est pas nécessaire de calculer tout l'ensemble des cumulants de \mathbf{v} pour calculer $\mathcal{Q}_{\mathbf{v}}(\mathbf{M})$.

D'après l'équation (2.15), on a $\mathbf{z}[\cdot] = \mathbf{U} \mathbf{s}[\cdot]$. En utilisant l'indépendance mutuelle des sources, et en notant $\mathcal{Q}_{\mathbf{z}}(\mathbf{M}) = \mathcal{Q}_{\mathbf{z}[\cdot]}(\mathbf{M})$, on peut montrer la relation suivante [CS93] :

$$\mathcal{Q}_{\mathbf{z}}(\mathbf{M}) = \mathbf{U} \mathbf{\Lambda}(\mathbf{M}) \mathbf{U}^T \quad (2.70)$$

avec

$$\mathbf{\Lambda}(\mathbf{M}) = \text{diag}(k_{s_1} \mathbf{u}_1^T \mathbf{M} \mathbf{u}_1, \dots, k_{s_1} \mathbf{u}_n^T \mathbf{M} \mathbf{u}_n) \quad (2.71)$$

où \mathbf{u}_i désigne la j^{eme} colonne de \mathbf{U} . Il apparaît donc que \mathbf{U} est une base orthonormale dans laquelle la matrice $\mathcal{Q}_{\mathbf{z}}(\mathbf{M})$ est diagonale, $\forall \mathbf{M}$. La matrice \mathbf{U} peut donc être estimée par diagonalisation d'une matrice de cumulants ayant des valeurs propres distinctes, aux indéterminations près de la diagonalisation, c-à-d, à des permutations de colonnes et inversion de signe près. Par exemple, si on choisit $\mathbf{M} = \mathbf{I}_n$ cela implique que les sources doivent avoir des kurtosis distincts². Le choix de \mathbf{M} est donc délicat.

En outre, si l'on choisit \mathbf{M} telle que les valeurs propres de $\mathcal{Q}_{\mathbf{z}}(\mathbf{M})$ ne sont pas bien séparées, une faible erreur dans l'estimation de $\mathcal{Q}_{\mathbf{z}}(\mathbf{M})$ peut induire de larges erreurs sur l'estimation de ses vecteurs propres.

Ces considérations ont amené Cardoso & Souloumiac à introduire le concept de *diagonalisation simultanée*. Soit $\mathbf{M} \in \mathbb{R}^{n \times n}$ une matrice quelconque, on définit la mesure de «diagonalité» suivante :

$$\text{off}(\mathbf{M}) = \sum_{i \neq j} m_{ij}^2 \quad (2.72)$$

²Encore une fois, on constate que si plus d'une source est gaussienne, alors $\mathbf{\Lambda}(\mathbf{M})$ possède deux termes diagonaux nuls, ce qui implique obligatoirement une indétermination sévère sur \mathbf{U} .

Soit \mathcal{M} une base de l'ensemble des matrices de $\mathbb{R}^{n \times n}$. Par exemple, si $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ désigne la base canonique de $\mathbb{R}^{n \times 1}$, on peut choisir $\mathcal{M} = \{\mathbf{e}_i \mathbf{e}_j^T, (i, j) \in \llbracket 1, n \rrbracket^2\}$. On a alors [Car99] :

$$\mathcal{C}_{Jade}^\circ\{\mathbf{V} \mathbf{z}\} = \sum_{\mathbf{M} \in \mathcal{M}} \text{off}(\mathbf{V}^T \mathcal{Q}_z(\mathbf{M}) \mathbf{V}) \quad (2.73)$$

On voit donc que la minimisation de la fonction de contraste $\mathcal{C}_{Jade}^\circ\{\mathbf{V} \mathbf{z}\}$ par rapport à la matrice orthonormale \mathbf{V} peut être réalisée par la diagonalisation simultanée de l'ensemble des matrices $\{\mathcal{Q}_z(\mathbf{M}_i), \mathbf{M}_i \in \mathcal{M}\}$. Le fait de choisir cet ensemble de matrices tel que \mathcal{M} soit une base de $\mathbb{R}^{n \times n}$ plutôt que de choisir \mathcal{M} arbitrairement permet de garantir (2.73), c-à-d, de garantir que la diagonalisation simultanée reste équivalente à la minimisation d'une fonction de contraste, et donc de garantir l'équivariance de la méthode [Car99].

D'autres choix de bases de $\mathbb{R}^{n \times n}$ peuvent être utilisés. Certains choix astucieux permettent de limiter le nombre de matrices à diagonaliser simultanément [Car99, CS93].

L'algorithme de diagonalisation simultanée utilisé dans JADE est inspiré de la méthode de Jacobi pour le calcul des éléments propres d'une matrice réelle symétrique [GL96]. L'algorithme, proposé par Cardoso & Souloumiac [CS96], permet de réaliser une diagonalisation simultanée *approchée* car en pratique les matrices de cumulants ne sont pas *rigoureusement* diagonalisables simultanément. Cet algorithme est décrit en annexe A.

Résumé de JADE :

- 1) Estimation d'une matrice de blanchiment $\hat{\mathbf{W}}$ d'après les éléments propres de $\hat{\mathcal{R}}_{\mathbf{x}\mathbf{x}}^a$,
- 2) Calculer $\mathbf{z} = \hat{\mathbf{W}} \mathbf{x}$,
- 3) Estimation de l'ensemble de matrices de cumulants $\{\hat{\mathcal{Q}}_z(\mathbf{M}_i), \mathbf{M}_i \in \mathcal{M}\}$: les cumulants sont estimés classiquement en remplaçant les espérances mathématiques dans les définitions par des moyennes temporelles (ce qui suppose que les signaux sont ergodiques),
- 4) Estimation de la matrice $\hat{\mathbf{U}}$ par diagonalisation simultanée de l'ensemble des matrices de cumulants,
- 5) Une estimation de \mathbf{A} est $\hat{\mathbf{A}} = \hat{\mathbf{W}}^\# \hat{\mathbf{U}}$, une estimation de $\mathbf{A}^\#$ est $\hat{\mathbf{B}} = \hat{\mathbf{U}}^H \hat{\mathbf{W}}$.

^aDans l'implémentation de JADE en contexte bruité [CS93], la variance du bruit est estimée puis prise en compte dans l'estimation de \mathbf{W} à la manière de la méthode d'orthonormalisation présentée au paragraphe 2.2.1

2.2 Séparation de sources stationnaires colorées (modèle \mathbb{M}_3)

On considère dans ce paragraphe la séparation de sources stationnaires colorées répondant au modèle \mathbb{M}_3 . Par rapport aux méthodes de séparation de sources i.i.d précédemment étudiées, cela suppose que les sources sont identiquement distribuées dans le temps mais qu'elles possèdent désormais une coloration : l'auto-covariance des séquences s_i ne se résume donc plus à une impulsion centrée sur $\tau = 0$. La stationnarité implique que la matrice de covariance des sources $\mathcal{R}_{\mathbf{ss}}[t, \tau]$ ne dépend que du retard τ . On écrira donc :

$$\mathcal{R}_{\mathbf{ss}}[t, \tau] = \mathcal{R}_{\mathbf{ss}}[\tau] \quad (2.74)$$

La plupart des signaux réels possèdent une telle coloration, qui est une information exploitable pour la séparation de sources. Comme évoqué au paragraphe 1.3.3, l'existence de cette coloration permet d'élaborer des méthodes se limitant à l'utilisation de statistiques d'ordre 2.

L'algorithme SOBI [BAMCM97] que nous allons présenter fonctionne sur le même principe que JADE : il basé sur la diagonalisation simultanée d'un ensemble de matrices correspondant à une représentation spatiale pertinente des séquences à séparer. Dans le cas de séquences i.i.d, JADE utilise des matrices de cumulants, dans le cadre de sources stationnaires colorées SOBI utilise des matrices de covariance. Comme JADE, SOBI procède en deux étapes : 1) Estimation d'une matrice d'orthonormalisation, 2) Estimation du facteur orthonormal manquant.

Nous décrivons SOBI dans le cadre général de séquences à valeurs éventuellement complexes et en contexte bruité.

2.2.1 SOBI en théorie

D'après $\mathbf{x} = \mathbf{A} \mathbf{s} + \mathbf{n}$ et avec l'hypothèse (H4) sur le bruit, on a :

$$\mathcal{R}_{\mathbf{xx}}[\tau] = \mathbf{A} \mathcal{R}_{\mathbf{ss}}[\tau] \mathbf{A}^H + \delta[\tau] \sigma^2 \mathbf{I}_n \quad (2.75)$$

où $\delta[\tau]$ désigne la fonction Delta définie par :

$$\begin{cases} \delta[0] = 1 \\ \delta[\tau] = 0 \quad \text{si } \tau \neq 0 \end{cases} \quad (2.76)$$

Orthonormalisation

La première étape de l'algorithme SOBI consiste à calculer une matrice d'orthonormalisation $\mathbf{W} \in \mathbb{C}^{n \times m}$ de rang plein telle que :

$$\mathbf{W} \mathbf{A} \mathcal{R}_{\mathbf{ss}}[0] \mathbf{A}^H \mathbf{W}^H = \mathbf{I}_n \quad (2.77)$$

Nous avons vu au paragraphe 1.2 comment il est possible, dans le cas non bruité, de calculer \mathbf{W} d'après les éléments propres de $\mathcal{R}_{\mathbf{xx}}[0] = \mathbf{A} \mathcal{R}_{\mathbf{ss}}[0] \mathbf{A}^H$. En milieu bruité, d'après l'équation (2.75) on a :

$$\mathcal{R}_{\mathbf{xx}}[0] = \mathbf{A} \mathcal{R}_{\mathbf{ss}}[0] \mathbf{A}^H + \sigma^2 \mathbf{I}_n \quad (2.78)$$

Cette expression de $\mathcal{R}_{\mathbf{xx}}[0]$ permet de calculer \mathbf{W} mais aussi σ^2 si m est *strictement* supérieur à n .

Soient $[\lambda_1, \dots, \lambda_m]$, les valeurs propres de $\mathcal{R}_{\mathbf{xx}}[0]$ rangées dans l'ordre décroissant. Soient $[\mathbf{h}_1, \dots, \mathbf{h}_m]$ des vecteurs propres normés correspondant. $\mathbf{A} \mathcal{R}_{\mathbf{ss}}[0] \mathbf{A}^H$ étant une matrice de rang n , les $m - n$ plus petites valeurs propres de $\mathcal{R}_{\mathbf{xx}}[0]$ sont donc égales à σ^2 et la matrice \mathbf{W} définie par

$$\mathbf{W} = \left[\frac{1}{\sqrt{\lambda_1 - \sigma^2}} \mathbf{h}_1 \dots \frac{1}{\sqrt{\lambda_n - \sigma^2}} \mathbf{h}_n \right]^H \quad (2.79)$$

vérifie l'équation (2.77).

Il est à noter qu'avec \mathbf{W} ainsi construite (c-à-d, avec des lignes orthogonales), les composantes de $\mathbf{z} = \mathbf{W} \mathbf{x}$ sont mutuellement décorréliées à $\tau = 0$. En effet, on a :

$$\mathcal{R}_{\mathbf{zz}}[0] = \mathbf{W} \mathcal{R}_{\mathbf{xx}}[0] \mathbf{W}^H \quad (2.80)$$

$$= \mathbf{W} (\mathbf{A} \mathcal{R}_{\mathbf{ss}}[0] \mathbf{A}^H) \mathbf{W}^H + \sigma^2 \mathbf{W} \mathbf{W}^H \quad (2.81)$$

$$= \mathbf{I}_n + \text{diag} \left(\frac{\sigma^2}{\lambda_1 - \sigma^2}, \dots, \frac{\sigma^2}{\lambda_n - \sigma^2} \right) \quad (2.82)$$

En revanche, dans le cas bruité, les composantes de \mathbf{z} ne sont plus de variance unitaire. On pourrait comme précédemment lever l'indétermination sur le gain des sources en fixant $\mathcal{R}_{\mathbf{ss}}[0] = \mathbf{I}_n$. Cependant, afin d'adopter une présentation cohérente avec l'ensemble des travaux présentés par la suite, nous choisissons d'intégrer $\mathcal{R}_{\mathbf{ss}}[0]$ aux calculs qui suivent.

Rotation

Ayant calculé une matrice \mathbf{W} vérifiant (2.77), d'après la propriété 3, il ne nous reste qu'à calculer $\mathbf{U} = \mathbf{W} \mathbf{A} \mathcal{R}_{ss}[0]^{\frac{1}{2}}$ pour obtenir \mathbf{A} ou $\mathbf{A}^\#$.

On définit $\forall \tau \neq 0$, les matrices suivantes :

$$\underline{\mathcal{R}}_{xx}[\tau] = \mathbf{W} \mathcal{R}_{xx}[\tau] \mathbf{W}^H \quad (2.83)$$

On a alors d'après (2.75), $\forall \tau \neq 0$:

$$\underline{\mathcal{R}}_{xx}[\tau] = \mathbf{W} \mathbf{A} \mathcal{R}_{ss}[\tau] \mathbf{A}^H \mathbf{W}^H \quad (2.84)$$

$$= (\mathbf{U} \mathcal{R}_{ss}[0]^{-\frac{1}{2}}) \mathcal{R}_{ss}[\tau] (\mathbf{U} \mathcal{R}_{ss}[0]^{-\frac{1}{2}})^H \quad (2.85)$$

$$= \mathbf{U} (\mathcal{R}_{ss}[0]^{-1} \mathcal{R}_{ss}[\tau]) \mathbf{U}^H \quad (2.86)$$

Les covariances des sources $\mathcal{R}_{ss}[\tau]$ étant des matrices diagonales *pour tout retard* τ , on en déduit que \mathbf{U} représente une base orthonormale dans laquelle $\underline{\mathcal{R}}_{xx}[\tau]$ est diagonale *pour tout retard* τ .

Si les sources répondent au modèle \mathbb{M}_3 alors il existe τ_1 tel que les termes de $\mathcal{R}_{ss}[0]^{-1} \mathcal{R}_{ss}[\tau_1]$ sont distincts, \mathbf{U} peut être estimée par diagonalisation de la matrice $\underline{\mathcal{R}}_{xx}[\tau_1]$. Cependant, comme pour JADE, il est en pratique recommandé de diagonaliser simultanément un ensemble de matrices $\{\underline{\mathcal{R}}_{xx}[\tau_i], i = 1, \dots, K\}$. L'intérêt de la diagonalisation simultanée de plusieurs matrices est de réduire les indéterminations possibles de la diagonalisation (voir modèle \mathbb{M}'_2) et de produire une estimation plus robuste de \mathbf{U} . En ce sens, SOBI est une extension de l'algorithme AMUSE [TLSH91] basé quant à lui sur la diagonalisation simultanée de deux matrices de covariances $\underline{\mathcal{R}}_{xx}[\tau]$. Une autre méthode similaire est présentée dans [Fét88].

La diagonalisation simultanée produit une matrice \mathbf{U}_{ds} telle que :

$$\mathbf{U}_{ds} = \mathbf{U} \mathbf{P} \Phi \quad (2.87)$$

où $\mathbf{P} \in \mathbb{R}^{n \times n}$ est une matrice de permutation et $\Phi \in \mathbb{C}^{n \times n}$ est une matrice diagonale à coefficients de modules unitaires. On alors :

$$\tilde{\mathbf{A}} \stackrel{\text{def}}{=} \mathbf{W}^\# \mathbf{U}_{ds} \quad (2.88)$$

$$= \mathbf{W}^\# \mathbf{U} \mathbf{P} \Phi \quad (2.89)$$

$$= \mathbf{A} \mathcal{R}_{ss}[0]^{\frac{1}{2}} \mathbf{P} \Phi \quad (2.90)$$

La matrice $\tilde{\mathbf{A}}$ est donc équivalente à \mathbf{A} .

2.2.2 SOBI en pratique

En supposant les signaux ergodiques, les matrices de covariances $\mathcal{R}_{xx}[\tau]$ sont estimées en pratique par :

$$\hat{\mathcal{R}}_{xx}[\tau] = \frac{1}{T - \tau} \sum_{t=0}^{T-\tau-1} \mathbf{x}[t + \tau] \mathbf{x}[t]^H \quad (2.91)$$

Résumé de SOBI :

- 1) Calcul de $\hat{\sigma}^2$ par moyenne des $m - n$ plus petites valeurs propres de $\hat{\mathcal{R}}_{\mathbf{xx}}[0]$ (si $m > n$),
- 2) Calcul de $\hat{\mathbf{W}}$ d'après $\hat{\sigma}^2$ et les éléments propres de $\hat{\mathcal{R}}_{\mathbf{xx}}[0]$,
- 3) Pour un ensemble de retards arbitraires $\{\tau_1, \dots, \tau_K\}$ on calcule :

$$\hat{\mathcal{R}}_{\mathbf{xx}}[\tau] = \hat{\mathbf{W}} \hat{\mathcal{R}}_{\mathbf{xx}}[\tau] \hat{\mathbf{W}}^H$$

- 4) Calcul de $\hat{\mathbf{U}}$ par diagonalisation simultanée de l'ensemble $\{\hat{\mathcal{R}}_{\mathbf{xx}}[\tau_i], i \in \llbracket 1, K \rrbracket\}$,
- 5) Une estimation de \mathbf{A} est $\hat{\mathbf{A}} = \hat{\mathbf{W}}^\# \hat{\mathbf{U}}$. Une estimation de $\mathbf{A}^\#$ est $\hat{\mathbf{B}} = \hat{\mathbf{U}}^H \hat{\mathbf{W}}$.

2.2.3 Discussion

On voit que l'existence d'une coloration temporelle des sources est une donnée permettant de mettre en oeuvre un algorithme de séparation simple et, comme nous le verrons dans les simulations de la partie 3.3.2, efficace. Dans la mesure où la plupart des signaux réels possèdent une coloration, SOBI est bien adapté à leur séparation. Remarquons qu'en pratique l'ensemble des retards $\{\tau_i, i = 1, \dots, K\}$ ne peut pas être connu à l'avance et qu'il devra être choisi arbitrairement. Toutefois, dans nos simulations, le choix et le nombre des retards ne nous a pas paru sensiblement affecter les performances de séparation de la méthode.

Enfin, remarquons que SOBI est équivalent à la minimisation d'une fonction de contraste orthonormale, ce qui lui confère la propriété d'équivariance [CL96]. En effet, SOBI correspond à la minimisation du contraste suivant :

$$C_{Sobi}^o\{\mathbf{V} \mathbf{z}\} = \sum_{i=1}^K \text{off}(\mathbf{V}^H \mathcal{R}_{\mathbf{zz}}[\tau_i] \mathbf{V}) \quad (2.92)$$

Chapitre 3

Séparation de sources non-stationnaires en mélange linéaire instantané

Nous nous intéressons dans ce chapitre à la séparation de sources non-stationnaires, c'est-à-dire dont les statistiques varient au cours du temps. Nous verrons dans ce chapitre que les méthodes de séparation présentées au chapitre précédent, et donc construites sur l'hypothèse de stationnarité des sources, peuvent donner des résultats acceptables sur des signaux non-stationnaires. Cependant, nous allons voir que les performances de séparation peuvent être accrues en *utilisant* le caractère non-stationnaire des sources.

Nous présentons essentiellement deux méthodes dans ce chapitre. La première méthode, proposée par Pham & Cardoso [PC01] et présentée dans la partie 3.1, est basée les hypothèses du modèle \mathbb{M}_2 . Les auteurs proposent une solution au sens du MV et montrent que la minimisation du critère résultant peut être réalisée par diagonalisation simultanée. Les auteurs utilisent l'hypothèse que la densité de chaque échantillon des sources $\mathbf{s}[\cdot]$ est gaussienne, ce qui signifie que seules les statistiques d'ordre 2 des sources sont exploitées.

La seconde méthode, que nous avons développée, présentée dans la partie 3.2, repose uniquement sur les hypothèses générales définissant le modèle \mathbb{M}_1 . Cette méthode est basée sur l'approximation du SWVS des sources par des Représentations Temps-Fréquence Spatiales de la classe de Cohen et sur la recherche d'*auto-termes simples* des sources.

3.1 Séparation de sources non-stationnaires blanches (modèle \mathbb{M}_2)

Une méthode simple...

Supposons que les sources soient modélisées par le modèle \mathbb{M}_2 , c-à-d par des séquences non-stationnaires blanches mutuellement décorréées. Comme nous l'avons vu précédemment, le fait que les sources soient blanches implique que $\mathcal{R}_{\mathbf{ss}}[t, \tau] = \mathbf{0}_n$ pour $\tau \neq 0$, et on note plus simplement :

$$\mathcal{R}_{\mathbf{ss}}[t, 0] = \mathcal{R}_{\mathbf{ss}}[t] = \text{diag}(r_{s_1 s_1}[t], \dots, r_{s_n s_n}[t]) \quad (3.1)$$

Dans le cas non-bruité, on a :

$$\mathcal{R}_{\mathbf{xx}}[t] = \mathbf{A} \mathcal{R}_{\mathbf{ss}}[t] \mathbf{A}^H \quad (3.2)$$

Disposant d'une matrice d'orthonormalisation \mathbf{W} , une méthode simple de séparation, inspirée de SOBI, consisterait à diagonaliser simultanément un ensemble de matrices $\underline{\mathcal{R}}_{\mathbf{xx}}[t] = \mathbf{W} \mathcal{R}_{\mathbf{xx}}[t] \mathbf{W}^H$ pour différents instants t_1, \dots, t_K . En supposant les observations \mathbf{x} stationnaires par morceaux, une estimation de $\mathcal{R}_{\mathbf{xx}}[t]$ pourrait être réalisée par une moyenne temporelle locale, approche

évoquée dans [PS00].

Pham & Cardoso [PC01] ont montré dans le cas non-bruité, réel et déterminé ($m = n$) que cette approche peut être connectée à l'estimateur du MV de \mathbf{A}^{-1} en modélisant les sources par des séquences gaussiennes stationnaires par morceaux (en plus des hypothèses de \mathbb{M}_2). Les auteurs montrent en outre que dans le cas $m = n$ un algorithme de diagonalisation de matrices hermitiennes positives permet d'estimer \mathbf{A} sans avoir recours à une orthonormalisation préalable.

Écriture du critère de MV

Supposons donc que les sources et la matrice de mélange sont à valeurs réelles avec $\mathbf{s} \in \mathbb{M}_2$, $\mathbf{n} = \mathbf{0}$ et $m = n$. En outre, la densité de chaque échantillon $\mathbf{s}[\cdot]$ des sources est supposée gaussienne, c-à-d :

$$\forall \mathbf{u} \in \mathbb{R}^{n \times 1}, \quad p_{\mathbf{s}[t]}(\mathbf{u}) = \frac{1}{\sqrt{\det(2\pi \mathcal{R}_{\mathbf{ss}}[t])}} \exp\left\{-\frac{1}{2} \mathbf{u}^T \mathcal{R}_{\mathbf{ss}}[t]^{-1} \mathbf{u}\right\} \quad (3.3)$$

Sous hypothèse gaussienne, les sources supposées blanches et mutuellement décorrélées deviennent indépendamment distribuées et mutuellement indépendantes. Ceci va nous permettre d'écrire simplement le critère du MV associé à toutes ces hypothèses.

D'après l'équation (3.3), on a :

$$\ln p_{\mathbf{s}[t]}(\mathbf{u}) = -\frac{1}{2} \ln \det(2\pi \mathcal{R}_{\mathbf{ss}}[t]) - \frac{1}{2} \mathbf{u}^T \mathcal{R}_{\mathbf{ss}}[t]^{-1} \mathbf{u} \quad (3.4)$$

$$= -\frac{1}{2} \ln \det(2\pi \mathcal{R}_{\mathbf{ss}}[t]) - \frac{1}{2} \text{trace}(\mathcal{R}_{\mathbf{ss}}[t]^{-1} \mathbf{u} \mathbf{u}^T) \quad (3.5)$$

Avec $\mathbf{x}[t] = \mathbf{A} \mathbf{s}[t]$, la densité de probabilité de $\mathbf{x}[t]$ est :

$$p_{\mathbf{x}[t]}(\mathbf{u}) = |\det(\mathbf{A})|^{-1} p_{\mathbf{s}[t]}(\mathbf{A}^{-1} \mathbf{u}) \quad (3.6)$$

D'où :

$$\ln p_{\mathbf{x}[t]}(\mathbf{u}) = -\ln \det(\mathbf{A}) + \ln p_{\mathbf{s}[t]}(\mathbf{A}^{-1} \mathbf{u}) \quad (3.7)$$

$$= -\ln \det(\mathbf{A}) - \frac{1}{2} \ln \det(2\pi \mathcal{R}_{\mathbf{ss}}[t]) - \frac{1}{2} \text{trace}(\mathcal{R}_{\mathbf{ss}}[t]^{-1} \mathbf{A}^{-1} \mathbf{u} \mathbf{u}^T \mathbf{A}^{-T}) \quad (3.8)$$

Les sources étant supposées indépendamment distribuées et mutuellement indépendantes, les observations sont également indépendamment distribuées, d'où :

$$p_{\mathbf{x}}(\mathbf{x}) = \prod_{t=0}^{T-1} p_{\mathbf{x}[t]}(\mathbf{x}[t]) \quad (3.9)$$

D'où :

$$\frac{1}{T} \ln p_{\mathbf{x}}(\mathbf{x}) = \frac{1}{T} \sum_{t=0}^{T-1} \ln p_{\mathbf{x}[t]}(\mathbf{x}[t]) \quad (3.10)$$

En posant $\mathbf{B} = \mathbf{A}^{-1}$, la log-vraisemblance de \mathbf{B} (et $\mathcal{R}_{\mathbf{ss}}[\cdot]$) par rapport à \mathbf{x} s'écrit donc :

$$L(\mathbf{B}, \mathcal{R}_{\mathbf{ss}}[\cdot]) = \ln \det(\mathbf{B}) - \frac{1}{2T} \sum_{t=0}^{T-1} \ln \det(2\pi \mathcal{R}_{\mathbf{ss}}[t]) + \text{trace}(\mathcal{R}_{\mathbf{ss}}[t]^{-1} \mathbf{B} \mathbf{x}[t] \mathbf{x}[t]^T \mathbf{B}^T)$$

En notant comme précédemment $\mathbf{y}[t] = \mathbf{B} \mathbf{x}[t]$, le gradient *relatif* de $L(\mathbf{B}, \mathcal{R}_{ss}[\cdot])$ par rapport à \mathbf{B} s'écrit [PC01] :

$$\nabla_{\mathbf{B}} L(\mathbf{B}, \mathcal{R}_{ss}[\cdot]) = \frac{1}{T} \sum_{t=0}^{T-1} \mathcal{R}_{ss}[t]^{-1} \mathbf{y}[t] \mathbf{y}[t]^T - \mathbf{I}_n \quad (3.11)$$

A $\mathcal{R}_{ss}[\cdot]$ fixé, les maxima de $L(\mathbf{B}, \mathcal{R}_{ss}[\cdot])$ sont donc caractérisés par :

$$\frac{1}{T} \sum_{t=0}^{T-1} \mathcal{R}_{ss}[t]^{-1} \mathbf{y}[t] \mathbf{y}[t]^T = \mathbf{I}_n \quad (3.12)$$

c-à-d :

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, \quad \frac{1}{T} \sum_{t=0}^{T-1} \frac{y_i[t] y_j[t]}{r_{s_i s_i}[t]} = \delta_{ij} \quad (3.13)$$

Cette expression possède une analogie évidente avec l'expression du gradient relatif de la vraisemblance de \mathbf{B} dans le cas de sources i.i.d (paragraphe 2.1.2). Les termes non-diagonaux ($i \neq j$) de cette équation matricielle expriment une forme de décorrélation «moyenne» entre les composantes de \mathbf{y} . Cependant nous n'avons plus ici besoin de réaliser la décorrélation non-linéaire des observations (équation (2.47)) : ici les statistiques d'ordre 2 suffisent, c'est la diversité temporelle qui est utilisée.

La résolution des équations données par (3.13) nécessite la connaissance *a priori* du profil de variance des sources $\mathcal{R}_{ss}[\cdot]$ qui n'est généralement évidemment pas disponible en pratique. Comme pour l'écriture du critère de MV dans le cas de sources i.i.d (paragraphe 2.1.2), nous allons voir que le critère $L(\mathbf{B}, \mathcal{R}_{ss}[\cdot])$ peut être maximisé suivant \mathbf{B} et $\mathcal{R}_{ss}[\cdot]$ et, en supposant les sources stationnaires par morceaux, ramené à la diagonalisation simultanée d'un ensemble de covariances empiriques.

Le modèle stationnaire par morceaux

Supposons que l'intervalle $\llbracket 0, T-1 \rrbracket$ puisse être divisé en K intervalles I_1, \dots, I_K sur lesquels les sources sont stationnaires, c-à-d, tels que, $\forall l \in \llbracket 1, K \rrbracket$:

$$\forall t \in I_l, \quad \mathcal{R}_{ss}[t] = \mathcal{R}_{ss}^{(l)} \quad (3.14)$$

où $\forall l \in \llbracket 1, K \rrbracket$, $\mathcal{R}_{ss}^{(l)}$ est une matrice constante. En outre, on pose :

$$\hat{\mathcal{R}}_{\mathbf{xx}}^{(l)} = \frac{1}{\#I_l} \sum_{t \in I_l} \mathbf{x}[t] \mathbf{x}[t]^T \quad (3.15)$$

où $\#I_l$ désigne le nombre d'éléments de I_l . On notera $w_l = \#I_l/T$ la proportion d'éléments dans I_l .

En remaniant l'expression de $L(\mathbf{B}, \mathcal{R}_{ss}[\cdot])$, nous allons montrer que le critère de MV peut s'écrire comme une somme de divergences de Kullback entre des gaussiennes de covariances $\mathcal{R}_{\mathbf{xx}}^{(l)}$ et $\hat{\mathcal{R}}_{\mathbf{xx}}^{(l)}$. D'après (3.8), et en utilisant pour des *matrices carrées* les relations $\det(\mathbf{M}\mathbf{N}) = \det(\mathbf{N}\mathbf{M})$ et $\det(\mathbf{M}^T) = \det(\mathbf{M})$, on a :

$$\begin{aligned} \ln p_{\mathbf{x}[t]}(\mathbf{u}) &= -\ln \det(\mathbf{A}) - \frac{1}{2} \ln \det(2\pi \mathcal{R}_{ss}[t]) - \frac{1}{2} \text{trace}(\mathcal{R}_{ss}[t]^{-1} \mathbf{A}^{-1} \mathbf{u} \mathbf{u}^T \mathbf{A}^{-T}) \\ &= -\frac{1}{2} [2 \ln \det(\mathbf{A}) + \ln \det(2\pi \mathcal{R}_{ss}[t]) + \text{trace}(\mathcal{R}_{ss}[t]^{-1} \mathbf{A}^{-1} \mathbf{u} \mathbf{u}^T \mathbf{A}^{-T})] \\ &= -\frac{1}{2} [\ln \det(\mathbf{A})^2 + \ln \det(\mathcal{R}_{ss}[t]) + n \ln 2\pi + \text{trace}(\mathcal{R}_{ss}[t]^{-1} \mathbf{A}^{-1} \mathbf{u} \mathbf{u}^T \mathbf{A}^{-T})] \\ &= -\frac{1}{2} [-\ln \det(\mathbf{A}^{-T} \mathcal{R}_{ss}[t]^{-1} \mathbf{A}^{-1}) + n \ln 2\pi + \text{trace}(\mathcal{R}_{ss}[t]^{-1} \mathbf{A}^{-1} \mathbf{u} \mathbf{u}^T \mathbf{A}^{-T})] \end{aligned}$$

On a donc d'après (3.10) :

$$\begin{aligned}
L(\mathbf{B}, \mathcal{R}_{\text{ss}}[\cdot]) &= \\
& - \frac{1}{2T} \sum_{l=1}^L \sum_{t \in I_l} \text{trace}(\mathcal{R}_{\text{ss}}[t]^{-1} \mathbf{B} \mathbf{x}[t] \mathbf{x}[t]^T \mathbf{B}^T) - \ln \det(\mathbf{B}^T \mathcal{R}_{\text{ss}}[t]^{-1} \mathbf{B}) + n \ln 2 \pi \\
& = - \frac{1}{2} \sum_{l=1}^L w_l \left[\text{trace}(\mathcal{R}_{\text{ss}}^{(l)-1} \mathbf{B} \hat{\mathcal{R}}_{\text{xx}}^{(l)} \mathbf{B}^T) - \ln \det(\mathbf{B}^T \mathcal{R}_{\text{ss}}^{(l)-1} \mathbf{B}) + n \ln 2 \pi \right] \\
& = \left\{ - \frac{1}{2} \sum_{l=1}^L w_l \left[\text{trace}(\mathcal{R}_{\text{ss}}^{(l)-1} \mathbf{B} \hat{\mathcal{R}}_{\text{xx}}^{(l)} \mathbf{B}^T) - \ln \det(\mathcal{R}_{\text{ss}}^{(l)-1} \mathbf{B} \hat{\mathcal{R}}_{\text{xx}}^{(l)} \mathbf{B}^T) \right] \right\} + \text{constante}
\end{aligned}$$

La divergence de Kullback $\mathcal{K}_g(\mathcal{R}_1 \parallel \mathcal{R}_2)$ entre deux vecteurs aléatoires centrés gaussiens de dimension n et de covariances \mathcal{R}_1 et \mathcal{R}_2 est donnée par [PC01] :

$$\mathcal{K}_g(\mathcal{R}_1 \parallel \mathcal{R}_2) = \frac{1}{2} [\text{trace}(\mathcal{R}_2^{-1} \mathcal{R}_1) - \ln \det(\mathcal{R}_2^{-1} \mathcal{R}_1) - n] \quad (3.16)$$

Le critère du MV s'écrit donc :

$$L(\mathbf{B}, \mathcal{R}_{\text{ss}}[\cdot]) \stackrel{\text{c}}{=} - \sum_{l=1}^L w_l \mathcal{K}_g(\mathbf{B} \hat{\mathcal{R}}_{\text{xx}}^{(l)} \mathbf{B}^T \parallel \mathcal{R}_{\text{ss}}^{(l)}) \quad (3.17)$$

De la même manière qu'au paragraphe 2.1.2, en utilisant la relation (2.52), si \mathcal{R} est une matrice positive et \mathcal{D} est une matrice positive diagonale alors :

$$\mathcal{K}_g(\mathcal{R} \parallel \mathcal{D}) = \mathcal{K}_g(\mathcal{R} \parallel \text{diag}(\mathcal{R})) + \mathcal{K}_g(\text{diag}(\mathcal{R}) \parallel \mathcal{D}) \quad (3.18)$$

D'où :

$$\begin{aligned}
L(\mathbf{B}, \mathcal{R}_{\text{ss}}[\cdot]) &\stackrel{\text{c}}{=} \\
& - \sum_{l=1}^L w_l \left\{ \mathcal{K}_g(\mathbf{B} \hat{\mathcal{R}}_{\text{xx}}^{(l)} \mathbf{B}^T \parallel \text{diag}(\mathbf{B} \hat{\mathcal{R}}_{\text{xx}}^{(l)} \mathbf{B}^T)) + \mathcal{K}_g(\text{diag}(\mathbf{B} \hat{\mathcal{R}}_{\text{xx}}^{(l)} \mathbf{B}^T) \parallel \mathcal{R}_{\text{ss}}^{(l)}) \right\} \quad (3.19)
\end{aligned}$$

Il alors est facile de maximiser $L(\mathbf{B}, \mathcal{R}_{\text{ss}}[\cdot])$ par rapport aux covariances des sources. Pour une valeur fixée de \mathbf{B} , il suffit de choisir $\forall l \in \llbracket 1, K \rrbracket$, $\mathcal{R}_{\text{ss}}^{(l)} = \text{diag}(\mathbf{B} \hat{\mathcal{R}}_{\text{xx}}^{(l)} \mathbf{B}^T)$. Le critère de vraisemblance ainsi maximisé par rapport aux covariances des sources devient :

$$L(\mathbf{B}) = - \sum_{l=1}^L w_l \mathcal{K}_g(\mathbf{B} \hat{\mathcal{R}}_{\text{xx}}^{(l)} \mathbf{B}^T \parallel \text{diag}(\mathbf{B} \hat{\mathcal{R}}_{\text{xx}}^{(l)} \mathbf{B}^T)) \quad (3.20)$$

En notant $\forall \mathbf{M} \in \mathbb{R}^{n \times n}$:

$$\text{off}_2(\mathbf{M}) \stackrel{\text{def}}{=} \mathcal{K}_g(\mathbf{M} \parallel \text{diag}(\mathbf{M})) \quad (3.21)$$

on a :

$$L(\mathbf{B}) = - \sum_{l=1}^L w_l \text{off}_2 \left\{ \mathbf{B} \hat{\mathcal{R}}_{\text{xx}}^{(l)} \mathbf{B}^T \right\} \quad (3.22)$$

La quantité off_2 est une mesure de diagonalité. En effet, d'après les propriétés de la divergence de Kullback, $\text{off}_2(\mathbf{M})$ est positif, nul ssi $\mathbf{M} = \text{diag}(\mathbf{M})$, c-à-d, si \mathbf{M} est diagonale. La maximisation du critère $L(\mathbf{B})$ est donc ramenée à la diagonalisation simultanée de l'ensemble de matrices positives hermitiennes $\{\hat{\mathcal{R}}_{\text{xx}}^{(l)}, l \in \llbracket 1, K \rrbracket\}$ au sens de la mesure de diagonalité off_2 .

La méthode exposée est nommée «BGML» dans [PC01] pour «Block-Gaussian Maximum Likelihood». La méthode est équivariante car elle peut s'exprimer en fonction du seul paramètre $\mathbf{y} = \mathbf{B}\mathbf{x}$ [CL96].

Un algorithme de diagonalisation simultanée de matrices carrées hermitiennes positives est décrit dans [Pha01]. Une implémentation en MATLAB de cet algorithme est disponible en ligne [Pha]. Cet algorithme dispense la méthode d'un blanchiment préalable et permet donc d'éviter les erreurs dues au blanchiment, qui ne sont pas rattrapables dans l'estimation du facteur unitaire. En outre, l'intérêt de la mesure off_2 est qu'elle ne dépend pas de l'échelle des matrices et qu'elle garde une connexion avec le maximum de vraisemblance.

En revanche Pham & Cardoso n'ont développé leur méthode que dans le cas $m = n$, l'extension au cas sur-déterminé est possible mais non triviale. En particulier la routine MATLAB de diagonalisation simultanée n'est disponible que pour le cas $m = n$.

Résumé de BGML :

- 1) Division de l'intervalle $\llbracket 0, T - 1 \rrbracket$ en K intervalles adjacents,
- 2) Calcul des K covariances locales $\{\hat{\mathcal{R}}_{\mathbf{xx}}^{(l)}, l \in \llbracket 1, K \rrbracket\}$,
- 3) Diagonalisation simultanée, *sans contrainte orthonormale*, de l'ensemble $\{\hat{\mathcal{R}}_{\mathbf{xx}}^{(l)}, l \in \llbracket 1, K \rrbracket\}$,
- 4) La diagonalisation simultanée fournit directement une estimation $\hat{\mathbf{B}}$ de \mathbf{A}^{-1} .

3.2 Contributions à une approche temps-fréquence (modèle \mathbb{M}_1)

Nous nous intéressons ici à une méthode de séparation basée sur les seules hypothèses du modèle \mathbb{M}_1 . Plutôt que d'ajouter des hypothèses au modèle \mathbb{M}_1 afin de faciliter l'estimation du SWVS, en supposant les sources stationnaires et colorées (modèle \mathbb{M}_3) ou non-stationnaires et blanches (modèle \mathbb{M}_2), nous allons nous intéresser à l'estimation directe du SWVS dans le cas général.

Le SWVS sera approximé par des Représentations Temps-Fréquence Spatiales (RTFS) de la classe de Cohen, et, à la manière de SOBI, la méthode que nous allons décrire consiste, après orthonormalisation, à diagonaliser simultanément un ensemble de matrices RTFS correspondant à plusieurs points temps-fréquence (t-f). L'idée d'utiliser des RTFS comme distribution conjointe des sources (à la place de matrices de covariances dans le cas stationnaire coloré ou de matrices de cumulants dans le cas i.i.d) a été introduite par Belouchrani et Amin dans [BA98]. Nous présentons deux contributions à leurs travaux :

- Les auteurs de [BA98] ne se placent pas explicitement dans un contexte *stochastique*. Nous proposons une nouvelle interprétation de leur méthode en termes stochastiques. Nous montrons que leur méthode peut être écrite en théorie sans aucune approximation au moyen du SWVS. Les RTFS s'interprètent alors simplement comme des estimations en pratique du SWVS.
- Contrairement à SOBI où le choix des retards peut être réalisé arbitrairement (en supposant qu'ils satisfont les conditions d'identifiabilité), nous allons voir que la sélection des points temps-fréquence correspondant aux matrices à diagonaliser simultanément est un point crucial de la méthode temps-fréquence. Nous proposons ainsi un nouveau critère de sélection basé sur l'identification de points t-f correspondant à des *auto-termes simples* des sources.

Nous présentons dans un premier temps la méthode théorique puis étudions dans un deuxième temps son implémentation en pratique.

3.2.1 Méthode théorique

Nous montrons dans ce paragraphe comment il est possible de calculer \mathbf{A} sans approximation, en théorie.

Du plan temps-retard au plan temps-fréquence

En environnement bruité et dans le contexte des hypothèses générales du modèle \mathbb{M}_1 , on a :

$$\mathcal{R}_{\mathbf{xx}}[t, \tau] = \mathbf{A} \mathcal{R}_{\mathbf{ss}}[t, \tau] \mathbf{A}^H + \delta[\tau] \sigma^2 \mathbf{I}_m \quad (3.23)$$

Soit encore, de manière équivalente :

$$\mathcal{S}_{\mathbf{xx}}[t, f] = \mathbf{A} \mathcal{S}_{\mathbf{ss}}[t, f] \mathbf{A}^H + 2 \sigma^2 \mathbf{I}_m \quad (3.24)$$

Comme pour JADE et SOBI, une méthode de séparation en deux étapes (orthonormalisation et rotation) peut être construite à partir de (3.24).

L'intérêt du passage au plan temps-fréquence réside dans l'existence d'un large panel d'outils t-f permettant l'estimation du SWVS.

Orthonormalisation et rotation

Soit comme au paragraphe 1.3.1 la matrice $\mathcal{R}_{\mathbf{ss}}$ définie par :

$$\mathcal{R}_{\mathbf{ss}} = \frac{1}{T} \sum_{t=0}^{T-1} \mathcal{R}_{\mathbf{ss}}[t, 0] = \text{diag}(r_{s1s1}, \dots, r_{s2s2}) \quad (3.25)$$

qui définit une «puissance temporelle moyenne» des sources. Soit \mathbf{W} une matrice d'orthonormalisation vérifiant :

$$\mathbf{W} \mathbf{A} \mathcal{R}_{\mathbf{ss}} \mathbf{A}^H \mathbf{W}^H = \mathbf{I}_n \quad (3.26)$$

On pose $\mathbf{U} = \mathbf{W} \mathbf{A} \mathcal{R}_{\mathbf{ss}}^{-\frac{1}{2}}$. On définit :

$$\underline{\mathcal{S}}_{\mathbf{xx}}[t, f] = \mathbf{W} (\mathcal{S}_{\mathbf{xx}}[t, f] - 2 \sigma^2 \mathbf{I}_m) \mathbf{W}^H \quad (3.27)$$

Avec les équations (3.24) et (3.27), on a :

$$\underline{\mathcal{S}}_{\mathbf{xx}}[t, f] = (\mathbf{W} \mathbf{A}) \mathcal{S}_{\mathbf{ss}}[t, f] (\mathbf{W} \mathbf{A})^H \quad (3.28)$$

$$= \mathbf{U} (\mathcal{R}_{\mathbf{ss}}^{-\frac{1}{2}} \mathcal{S}_{\mathbf{ss}}[t, f] \mathcal{R}_{\mathbf{ss}}^{-\frac{H}{2}}) \mathbf{U}^H \quad (3.29)$$

$$= \mathbf{U} (\mathcal{R}_{\mathbf{ss}}^{-1} \mathcal{S}_{\mathbf{ss}}[t, f]) \mathbf{U}^H \quad (3.30)$$

Les matrices $\mathcal{R}_{\mathbf{ss}}$ et $\mathcal{S}_{\mathbf{ss}}[t, f]$ étant diagonales, \mathbf{U} peut être obtenue par diagonalisation d'une matrice $\underline{\mathcal{S}}_{\mathbf{xx}}[t_1, f_1]$ arbitraire dont les valeurs propres sont distinctes (modèle \mathbb{M}_1). Comme précédemment, \mathbf{U} peut être plus généralement obtenue par diagonalisation simultanée d'un ensemble de matrices $\{\underline{\mathcal{S}}_{\mathbf{xx}}[t_i, f_i], i = 1, \dots, K\}$ (modèle \mathbb{M}'_1). La diagonalisation simultanée produit une matrice \mathbf{U}_{ds} telle que :

$$\mathbf{U}_{ds} = \mathbf{U} \mathbf{P} \mathbf{\Phi} \quad (3.31)$$

où $\mathbf{P} \in \mathbb{R}^{n \times n}$ est une matrice de permutation et $\mathbf{\Phi} \in \mathbb{C}^{n \times n}$ est une matrice diagonale à coefficients de modules unitaires. On a alors :

$$\tilde{\mathbf{A}} \stackrel{\text{def}}{=} \mathbf{W}^\# \mathbf{U}_{ds} \quad (3.32)$$

$$= \mathbf{W}^\# \mathbf{U} \mathbf{P} \mathbf{\Phi} \quad (3.33)$$

$$= \mathbf{A} \mathcal{R}_{\mathbf{ss}}[0]^{\frac{1}{2}} \mathbf{P} \mathbf{\Phi} \quad (3.34)$$

La matrice $\tilde{\mathbf{A}}$ est donc équivalente à \mathbf{A} .

Ainsi, disposant de $\mathcal{S}_{\mathbf{x}\mathbf{x}}[t, f]$, \mathbf{W} et σ^2 , il est possible de calculer une matrice équivalente à \mathbf{A} . Une approche similaire utilisant le Spectre Évolutif Spatial au lieu du SWVS est présentée dans [KA00]. Cependant, l'estimation en pratique du Spectre Évolutif requiert en général que les signaux soient localement stationnaires [SJ95]¹.

Enfin, la méthode t-f n'étant qu'une généralisation de SOBI, elle possède aussi la propriété d'équivariance.

3.2.2 Implémentation pratique

Nous montrons dans ce paragraphe comment estimer les quantités \mathbf{W} , σ^2 et $\mathcal{S}_{\mathbf{x}\mathbf{x}}[t, f]$ d'après l'unique réalisation des observations dont nous disposons.

Estimation de \mathbf{W} et σ^2

Avec :

$$\mathcal{R}_{\mathbf{x}\mathbf{x}} = \frac{1}{T} \sum_{t=0}^{T-1} \mathcal{R}_{\mathbf{x}\mathbf{x}}[t, 0] \quad (3.35)$$

On a :

$$\mathcal{R}_{\mathbf{x}\mathbf{x}} = \mathbf{A} \mathcal{R}_{\mathbf{s}\mathbf{s}} \mathbf{A}^H + \sigma^2 \mathbf{I}_m \quad (3.36)$$

\mathbf{W} et σ^2 peuvent donc être calculés en théorie d'après les éléments propres de $\mathcal{R}_{\mathbf{x}\mathbf{x}}$ de la même façon qu'au paragraphe 1.2. Le problème est que dans le cas non-stationnaire, l'estimation de $\mathcal{R}_{\mathbf{x}\mathbf{x}}$ n'est pas simple car elle nécessite l'estimation de $\mathcal{R}_{\mathbf{x}\mathbf{x}}[t, 0]$ à chaque instant t . En supposant les signaux localement stationnaires $\mathcal{R}_{\mathbf{x}\mathbf{x}}[t, 0]$ pourrait être estimé par moyenne temporelle locale. Cependant, l'estimation «brute» de $\mathcal{R}_{\mathbf{x}\mathbf{x}}[t, 0]$ par $\mathbf{x}[t] \mathbf{x}[t]^H$ s'est révélée être satisfaisante dans nos simulations. Avec les hypothèses sur le bruit, on a alors :

$$\hat{\mathcal{R}}_{\mathbf{x}\mathbf{x}} = \frac{1}{T} \sum_{t=0}^{T-1} \mathbf{x}[t] \mathbf{x}[t]^H \quad (3.37)$$

$$= \frac{1}{T} \sum_{t=0}^{T-1} \{ \mathbf{A} \mathbf{s}[t] \mathbf{s}[t]^H \mathbf{A}^H + \mathbf{n}[t] \mathbf{n}[t]^H + \mathbf{A} \mathbf{s}[t] \mathbf{n}[t]^H + \mathbf{n}[t] \mathbf{s}[t]^H \mathbf{A}^H \} \quad (3.38)$$

$$\approx \mathbf{A} \hat{\mathcal{R}}_{\mathbf{s}\mathbf{s}} \mathbf{A}^H + \sigma^2 \mathbf{I}_m \quad (3.39)$$

avec :

$$\hat{\mathcal{R}}_{\mathbf{s}\mathbf{s}} = \frac{1}{T} \sum_{t=0}^{T-1} \mathbf{s}[t] \mathbf{s}[t]^H \quad (3.40)$$

L'utilisation de cette approximation pour effectuer le blanchiment revient donc en fait à supposer que les *réalisations des sources* sont décorréliées au sens où la matrice $\hat{\mathcal{R}}_{\mathbf{s}\mathbf{s}}$ doit être idéalement diagonale. Ceci peut apparaître comme une sévère limitation, dans le cas $m = n$ et non bruité, par rapport à la méthode décrite au paragraphe 3.1 qui se dispense de blanchiment par l'utilisation d'un algorithme de diagonalisation simultanée de matrices hermitiennes positives. Cependant, les résultats présentés au paragraphe 3.3 montrent que la méthode basée sur les RTFS offre néanmoins de meilleures performances.

¹Les avantages du SWV sur le Spectre Évolutif sont décrits dans [SJ95].

Estimation du SWVS

Le SWVS peut s'écrire de la façon suivante :

$$\mathcal{S}_{\mathbf{x}\mathbf{x}}[t, f] = \mathbb{E}\{\hat{\mathcal{S}}_{\mathbf{x}\mathbf{x}}^{WV}[t, f]\} \quad (3.41)$$

où $\hat{\mathcal{S}}_{\mathbf{x}\mathbf{x}}^{WV}[t, f]$ désigne un échantillonnage fréquentiel arbitraire de la distribution de Wigner-Ville Spatiale définie pour des signaux à temps discret $\forall(t, f) \in \mathbb{Z} \times [-\frac{1}{2}, \frac{1}{2}]$ par :

$$\hat{\mathcal{S}}_{\mathbf{x}\mathbf{x}}^{WV}[t, f] = 2 \sum_{\tau} \mathbf{x}[t + \tau] \mathbf{x}[t - \tau]^H e^{-j4\pi f\tau} \quad (3.42)$$

Ainsi, la distribution de Wigner-Ville Spatiale est une estimation brute du SWVS basée sur la seule réalisation disponible des observations. Il est montré dans [Fla99, SJ95] que le lissage en temps et en fréquence de la distribution de Wigner-Ville, c-à-d, l'utilisation des RTF de la classe de Cohen, produit une meilleure estimation du SWV. Ainsi, pour un noyau de lissage $\phi[t, f]$ donné, on notera $\hat{\mathcal{S}}_{\mathbf{x}\mathbf{x}}^{\phi}[t, f]$ la RTFS définie par :

$$\hat{\mathcal{S}}_{\mathbf{x}\mathbf{x}}^{\phi}[t, f] = \sum_u \sum_v \phi[u - t, v - f] \hat{\mathcal{S}}_{\mathbf{x}\mathbf{x}}^{WV}[u, v] \quad (3.43)$$

De même que le SWVS, $\hat{\mathcal{S}}_{\mathbf{x}\mathbf{x}}^{\phi}[t, f]$ est pour chaque point (t, f) une matrice de dimensions $m \times m$. Ses termes diagonaux correspondent aux auto-RTF $\hat{\mathcal{S}}_{x_i x_i}^{\phi}[t, f]$ de chaque observation tandis que ses termes non-diagonaux correspondent aux RTF croisées de deux observations $\hat{\mathcal{S}}_{x_i x_j}^{\phi}[t, f]$. Les termes diagonaux de $\hat{\mathcal{S}}_{\mathbf{x}\mathbf{x}}^{\phi}[t, f]$ sont donc des *auto-termes* tandis que les termes non-diagonaux sont des *termes croisés*. Par la suite on suppose que le noyau est d'énergie unitaire, c-à-d :

$$\sum_u \sum_v \phi[u, v] = 1 \quad (3.44)$$

Ceci garantit que les RTF de la classe de Cohen, la distribution de Wigner-Ville et le SWV possèdent la même échelle d'amplitude. Le choix du noyau ϕ sera étudié dans la partie 3.3.

Disposant d'estimations $\hat{\mathbf{W}}$ et $\hat{\sigma}^2$, et utilisant $\hat{\mathcal{S}}_{\mathbf{x}\mathbf{x}}^{\phi}[t, f]$ comme estimation de $\mathcal{S}_{\mathbf{x}\mathbf{x}}[t, f]$, l'équation (3.30) devient :

$$\underline{\hat{\mathcal{S}}}_{\mathbf{x}\mathbf{x}}^{\phi}[t, f] \approx \mathbf{U} (\hat{\mathcal{R}}_{\mathbf{s}\mathbf{s}}^{-1} \hat{\mathcal{S}}_{\mathbf{s}\mathbf{s}}^{\phi}[t, f]) \mathbf{U}^H \quad (3.45)$$

avec :

$$\underline{\hat{\mathcal{S}}}_{\mathbf{x}\mathbf{x}}^{\phi}[t, f] \stackrel{\text{def}}{=} \hat{\mathbf{W}} (\hat{\mathcal{S}}_{\mathbf{x}\mathbf{x}}^{\phi}[t, f] - 2\hat{\sigma}^2 \mathbf{I}_m) \hat{\mathbf{W}}^H \quad (3.46)$$

A ce niveau, $\underline{\hat{\mathcal{S}}}_{\mathbf{x}\mathbf{x}}^{\phi}[t, f]$ est supposé connu, et la suite de nos travaux est dédiée à l'estimation de \mathbf{U} par une procédure de diagonalisation simultanée.

3.2.3 Sélection des points temps-fréquence

A la manière de SOBI, nous allons estimer \mathbf{U} par diagonalisation simultanée d'un ensemble de matrices $\{\underline{\hat{\mathcal{S}}}_{\mathbf{x}\mathbf{x}}^{\phi}[t_i, f_i], i \in \llbracket 1, K \rrbracket\}$. Si l'estimation ergodique classique des matrices de covariance s'est avérée suffisante dans les méthodes précédentes, un problème survient dans le cas temps-fréquence. $\hat{\mathcal{S}}_{\mathbf{s}\mathbf{s}}^{\phi}[t, f]$ n'étant qu'une estimation de $\mathcal{S}_{\mathbf{s}\mathbf{s}}[t, f]$, il apparaît en pratique que $\hat{\mathcal{S}}_{\mathbf{s}\mathbf{s}}^{\phi}[t, f]$ n'est pas diagonale pour n'importe quel point (t, f) , bien que théoriquement $\mathcal{S}_{\mathbf{s}\mathbf{s}}[t, f]$ le soit. Ainsi, nous ne pouvons pas diagonaliser n'importe quelle matrice $\underline{\hat{\mathcal{S}}}_{\mathbf{x}\mathbf{x}}^{\phi}[t, f]$ correspondant à n'importe quel point (t, f) . Avant de diagonaliser simultanément l'ensemble $\{\underline{\hat{\mathcal{S}}}_{\mathbf{x}\mathbf{x}}^{\phi}[t_i, f_i], i \in \llbracket 1, K \rrbracket\}$, il nous faut identifier *en aveugle* (c-à-d, à partir des observations seulement), un ensemble de points $\{(t_i, f_i), i \in \llbracket 1, K \rrbracket\}$ pour lesquels $\hat{\mathcal{S}}_{\mathbf{s}\mathbf{s}}^{\phi}[t, f]$ est véritablement diagonale. Cette sélection est une étape importante et non-triviale, qui conditionne les performances de la méthode.

Sélection des points de plus haute énergie

Dans l'article de référence [BA98], il est suggéré que les matrices $\hat{\mathcal{S}}_{\mathbf{xx}}^\phi[t, f]$ à diagonaliser simultanément soient choisies comme étant celles «de plus haute énergie dans le plan temps-fréquence». Les auteurs affirment que, puisque «les termes non-diagonaux de $\hat{\mathcal{S}}_{\mathbf{ss}}^\phi[t, f]$ sont des termes croisés, alors cette matrice est diagonale pour chaque point t-f correspondant à une concentration d'énergie». Toutefois, aucun critère de sélection explicite n'est proposé.

Sur une idée similaire, les mêmes auteurs proposent dans [BAMAZ01] le critère de sélection suivant :

$$\text{Sélectionner } (t, f) \text{ si } C_1[t, f] = \frac{\text{trace}(\hat{\mathcal{S}}_{\mathbf{xx}}^\phi[t, f])}{\|\hat{\mathcal{S}}_{\mathbf{xx}}^\phi[t, f]\|_F} \geq \varepsilon_{\text{Energ}} \quad (3.47)$$

où $\varepsilon_{\text{Energ}}$ désigne un seuil arbitraire. Les auteurs utilisent simplement les relations suivantes (avec l'équation 3.45 et \mathbf{U} étant orthonormale) :

$$\text{trace}(\hat{\mathcal{S}}_{\mathbf{xx}}^\phi[t, f]) \approx \text{trace}(\hat{\mathcal{R}}_{\mathbf{ss}}^{-1} \hat{\mathcal{S}}_{\mathbf{ss}}^\phi[t, f]) \quad (3.48)$$

$$\|\hat{\mathcal{S}}_{\mathbf{xx}}^\phi[t, f]\|_F \approx \|\hat{\mathcal{R}}_{\mathbf{ss}}^{-1} \hat{\mathcal{S}}_{\mathbf{ss}}^\phi[t, f]\|_F \quad (3.49)$$

qui permettent d'accéder à la trace et à la norme des RTFS des sources (normalisées par $\hat{\mathcal{R}}_{\mathbf{ss}}^{-1}$) d'après les observations.

Dans [BAMAZ01] les auteurs proposent en outre un critère de sélection de points t-f correspondant à des matrices «anti-diagonales» et proposent un algorithme d'«anti-diagonalisation» simultanée. Cette perspective n'est cependant pas étudiée ici (voir aussi [GTMA02] à ce sujet).

Sélection d'auto-termes simples

La méthode de sélection précédente nécessite le réglage du seuil $\varepsilon_{\text{Energ}}$. Ceci s'avère être difficile en pratique car aucun principe heuristique ne permet de guider le choix de la valeur de $\varepsilon_{\text{Energ}}$. Nous verrons en outre dans les résultats du paragraphe 3.3 que de faibles variations de la valeur $\varepsilon_{\text{Energ}}$ modifient sensiblement les performances de la méthode de séparation.

Nous proposons un nouveau critère de sélection plus robuste basé sur l'identification de matrices $\hat{\mathcal{S}}_{\mathbf{ss}}^\phi[t, f]$ diagonales *avec un seul terme non-nul sur la diagonale*. Cette idée vient simplement du fait qu'en pratique nous avons constaté que les matrices $\hat{\mathcal{S}}_{\mathbf{ss}}^\phi[t, f]$ diagonales n'ont *effectivement* qu'un seul terme non-nul sur la diagonale.

Ce phénomène s'explique intuitivement de la manière suivante : *si deux sources s_i et s_j possèdent une même fréquence f_0 au même instant t_0 , alors la RTF croisée entre ces deux sources au point (t_0, f_0) est vraisemblablement non nulle*. En d'autres termes, si $\hat{s}_{s_i s_i}^\phi(t_0, f_0)$ et $\hat{s}_{s_j s_j}^\phi(t_0, f_0)$ sont non nuls, il est vraisemblable que $\hat{s}_{s_i s_j}^\phi(t_0, f_0)$ et $\hat{s}_{s_j s_i}^\phi(t_0, f_0)$ soient aussi non nuls. Donc, si $\hat{\mathcal{S}}_{\mathbf{ss}}^\phi[t_0, f_0]$ est diagonale, il est vraisemblable qu'elle ne possède qu'un terme non-nul.

Par la suite, un point (t, f) correspondant à une matrice $\hat{\mathcal{S}}_{\mathbf{ss}}^\phi[t, f]$ ne possédant qu'un seul terme non-nul, situé sur la diagonale, sera désigné comme un point associé à un *auto-terme simple* des sources. Plus précisément, si $\hat{\mathcal{S}}_{\mathbf{ss}}^\phi[t, f]$ est diagonale avec un seul terme non-nul situé à la i^{eme} position, alors (t, f) sera un point t-f associé à un *auto-terme simple* de la i^{eme} source.

Nous proposons maintenant un critère permettant d'identifier les points t-f associés à des auto-termes simples des sources [FD04, HFDZ02].

Soit (t_0, f_0) un point t-f associé à un auto-terme simple des sources. Par définition, $\hat{\mathcal{S}}_{\mathbf{ss}}^\phi[t_0, f_0]$ est donc diagonale avec un seul terme non nul. $\hat{\mathcal{S}}_{\mathbf{ss}}^\phi[t_0, f_0]$ ne possède donc évidemment qu'une seule valeur propre non nulle. Avec l'équation (3.45) et \mathbf{U} orthonormale, on a :

$$\text{eig}(\hat{\mathcal{R}}_{\mathbf{ss}}^{-1} \hat{\mathcal{S}}_{\mathbf{ss}}^\phi[t_0, f_0]) \approx \text{eig}(\hat{\mathcal{S}}_{\mathbf{xx}}^\phi[t_0, f_0]) \quad (3.50)$$

où $\text{eig}(\mathbf{M})$ désigne l'ensemble des valeurs propres d'une matrice \mathbf{M} . Ainsi, $\hat{\mathcal{R}}_{\text{ss}}$ étant diagonale, une seule des valeurs propres de $\hat{\mathcal{S}}_{\text{xx}}^\phi[t_0, f_0]$ est non nulle, ce qui signifie en particulier que la plus grande des valeurs absolues des valeurs propres de $\hat{\mathcal{S}}_{\text{xx}}^\phi[t_0, f_0]$ égale la somme des valeurs absolues des valeurs propres de $\hat{\mathcal{S}}_{\text{xx}}^\phi[t_0, f_0]$.

Finalement, les point t-f associés à des auto-termes simples des sources vérifient :

$$C_2[t, f] \stackrel{\text{def}}{=} \frac{\max |\text{eig}(\hat{\mathcal{S}}_{\text{xx}}^\phi[t, f])|}{\sum |\text{eig}(\hat{\mathcal{S}}_{\text{xx}}^\phi[t, f])|} \quad (3.51)$$

$$\approx 1 \quad (3.52)$$

où la notation $|\text{eig}(\mathbf{M})|$ désigne l'ensemble des valeurs absolues des valeurs propres de \mathbf{M} .

La recherche d'auto-termes simples des sources est équivalente à la sélection des matrices $\hat{\mathcal{S}}_{\text{xx}}^\phi[t, f]$ de rang 1. Ceci signifie que tout point (t_0, f_0) tel que $C_2[t_0, f_0] = 1$ correspond à une matrice $\hat{\mathcal{S}}_{\text{ss}}^\phi[t, f]$ de rang 1, qui en théorie ne correspond pas obligatoirement à un auto-terme simple. Il est donc possible de sélectionner des matrices $\hat{\mathcal{S}}_{\text{xx}}^\phi[t, f]$ «pathologiques», qui correspondent à des matrices $\hat{\mathcal{S}}_{\text{ss}}^\phi[t, f]$ de rang 1 mais non-diagonales. Cependant, nous avons constaté qu'en pratique de tels cas sont rares, et lorsqu'ils surviennent, leur effet est réduit par le processus de diagonalisation simultanée de plusieurs matrices. Ces matrices «pathologiques» peuvent éventuellement être écartées par une méthode de diagonalisation simultanée *itérative* comme celle décrite dans [HFDZ02].

En pratique, $C_2[t, f]$ n'est calculé que pour les points t-f correspondant à une énergie significative des sources : afin d'éviter les matrices $\hat{\mathcal{S}}_{\text{ss}}^\phi[t, f]$ de coefficients tous proches de zéro, avec l'équation (3.45) on calcule $C_2[t, f]$ uniquement pour les points t-f tels que

$$\text{trace}(\hat{\mathcal{S}}_{\text{xx}}^\phi[t, f]) \geq \text{moy}(\text{trace}(\hat{\mathcal{S}}_{\text{xx}}^\phi[t, f])) \quad (3.53)$$

où $\text{moy}(\text{trace}(\hat{\mathcal{S}}_{\text{xx}}^\phi[t, f]))$ désigne la valeur moyenne de $\text{trace}(\hat{\mathcal{S}}_{\text{xx}}^\phi[t, f])$ sur tout le plan t-f.

Toutefois, pratiquement aucun point (t, f) ne satisfait $C_2[t, f] = 1$ car les matrices $\hat{\mathcal{S}}_{\text{ss}}^\phi[t, f]$ correspondant à des auto-termes simples des sources ne sont jamais strictement diagonales mais seulement quasi-diagonales. Une première idée consiste donc à utiliser le critère de sélection suivant :

$$\text{Sélectionner } (t, f) \text{ si } C_2[t, f] \geq 1 - \varepsilon \quad (3.54)$$

où ε est proche de zéro.

Afin d'éviter le problème du choix de ε nous proposons simplement de rechercher des points t-f pour lesquels $C_2[t, f]$ est élevé (proche de 1). Une solution satisfaisante consiste à chercher les maxima locaux de $C_2[t, f]$ sur l'ensemble du plan temps-fréquence. Par définition de $C_2[t, f]$, ses maxima locaux proches de 1 correspondent à des auto-termes simples de sources *quelconques*. Donc, ces maxima locaux sont distribués principalement le long des signatures t-f des sources (sauf là où les signatures t-f de plusieurs sources se chevauchent car dans ces régions au moins deux termes diagonaux de $\hat{\mathcal{S}}_{\text{ss}}^\phi[t, f]$ sont non nuls). Ainsi, *la recherche des maxima locaux de $C_2[t, f]$ assure la sélection de points t-f associés à des auto-termes simples de toutes les sources (ce qui est une condition nécessaire à l'unicité de la diagonalisation simultanée), et associés aux matrices les «plus diagonales» (ce qui améliore l'estimation de \mathbf{U})*. Ainsi, nous proposons de sélectionner les points t-f vérifiant :

$$\begin{cases} \|\nabla C_2[t, f]\|_2 & \leq \varepsilon_{\text{Grad}} \\ \nabla^2 C_2[t, f] & < 0 \end{cases} \quad (3.55)$$

où

$$\nabla C_2 = \left[\frac{\delta C_2}{\delta t}, \frac{\delta C_2}{\delta f} \right]^T$$

désigne le gradient² de C_2 et où

$$\nabla^2 C_2 = \begin{bmatrix} \frac{\delta^2 C}{\delta t^2} & \frac{\delta^2 C}{\delta t \delta f} \\ \frac{\delta^2 C}{\delta f \delta t} & \frac{\delta^2 C}{\delta f^2} \end{bmatrix}$$

désigne le Hessien³ de $C_2[t, f]$. Le seuil ε_{Grad} règle le nombre de points sélectionnés au voisinage d'un maximum local.

Nous donnons le nom de «TFBSS» à la méthode temps-fréquence utilisant le critère de sélection basé sur l'identification d'auto-termes simples des sources. Une implémentation MATLAB de TFBSS est disponible sur Internet sur [Fév] ainsi que sur «ICA Central» [Cen].

Résumé de TFBSS :

- 1) Calcul de $\hat{\sigma}^2$ par moyenne des $m - n$ plus petites valeurs propres de $\hat{\mathcal{R}}_{\mathbf{xx}} = \frac{1}{T} \sum \mathbf{x}[t] \mathbf{x}[t]^H$ (si $m > n$),
- 2) Calcul de $\hat{\mathbf{W}}$ d'après $\hat{\sigma}^2$ et les éléments propres de $\hat{\mathcal{R}}_{\mathbf{xx}}$,
- 3) Calcul de $\hat{\underline{\mathcal{S}}}_{\mathbf{xx}}^\phi[t, f]$ pour $(t, f) \in I_T \times I_F$, où I_T et I_F désignent des ensembles arbitraires d'index temporels et fréquentiels^a,
- 4) Calcul de $C_2[t, f]$ pour $(t, f) \in I_T \times I_F$,
- 5) Identification d'un ensemble de points $\{(t_i, f_i), i \in \llbracket 1, K \rrbracket\}$ associés à des auto-termes simples des sources (par méthode de gradient),
- 6) Calcul de $\hat{\mathbf{U}}$ par diagonalisation simultanée de l'ensemble $\{\hat{\underline{\mathcal{S}}}_{\mathbf{xx}}^\phi[t_i, f_i], i \in \llbracket 1, K \rrbracket\}$,
- 7) Une estimation de \mathbf{A} est $\hat{\mathbf{A}} = \hat{\mathbf{W}}^\# \hat{\mathbf{U}}$. Une estimation de $\mathbf{A}^\#$ est $\hat{\mathbf{B}} = \hat{\mathbf{U}}^H \hat{\mathbf{W}}$.

^aEn pratique une transformée de Hilbert pourra éventuellement être appliquée aux observations avant le calcul de leur RTFS. Ceci permet d'éviter le phénomène de repliement spectral [Fla99] et peut faciliter l'identification des auto-termes simples des sources.

Discussion

L'idée sous-jacente à la recherche d'auto-termes simples des sources est la localisation de points t-f où *une seule source est active*. Cette idée est commune à plusieurs travaux sur la séparation de mélanges linéaires instantanés, notamment sous-déterminés, basés sur des RTF *atomiques*. Les RTF *atomiques* sont des représentations linéaires alors que les RTF *énergétiques* que nous avons utilisées sont bilinéaires.

Par exemple, Abrard, Deville & White proposent une méthode simple d'identification de la matrice d'un mélange linéaire instantané éventuellement sous-déterminé [ADW01]. Leur méthode est basée sur la localisation de points t-f pour lesquels les Transformées de Fourier à Court-Terme (TFCT) de toutes les sources sauf une sont nulles. En effet, en supposant $m = n = 2$, en notant $\mathcal{F}_x[t, f]$ la TFCT d'un signal x pour une fenêtre de lissage temporel quelconque, on a :

$$F[t, f] \stackrel{\text{def}}{=} \frac{\mathcal{F}_{x_1}[t, f]}{\mathcal{F}_{x_2}[t, f]} = \frac{a_{11} \mathcal{F}_{s_1}[t, f] + a_{12} \mathcal{F}_{s_2}[t, f]}{a_{21} \mathcal{F}_{s_1}[t, f] + a_{22} \mathcal{F}_{s_2}[t, f]} \quad (3.56)$$

Si au point (t_0, f_0) seule la source s_1 est «active», c-à-d si $\mathcal{F}_{s_2}[t_0, f_0] = 0$, alors

$$F[t_0, f_0] = \frac{\mathcal{F}_{x_1}[t_0, f_0]}{\mathcal{F}_{x_2}[t_0, f_0]} = \frac{a_{11}}{a_{21}} \quad (3.57)$$

²Puisque C_2 est discret en temps et en fréquence, le gradient désigne simplement les différences de C_2 d'un échantillon à l'autre sur l'axe des temps ou des fréquences.

³Même remarque que pour le gradient : dans le cas discret le Hessien désigne simplement les différences des gradients d'un échantillon à l'autre.

La localisation du point (t_0, f_0) permet donc de calculer le rapport a_{11}/a_{21} . De même la localisation d'un point t-f où seule la source s_2 est active permet de calculer le rapport a_{12}/a_{22} , et donc d'identifier la matrice de mélange aux indéterminations près de la SAS. Les régions où seule une source est active sont déterminées comme étant celles sur lesquelles $F[t, f]$ est quasiment constant.

Si les sources ont des supports temps-fréquence disjoints, il alors envisageable de reconstruire les signaux sources après avoir localisé tout leur support t-f. Ceci est l'objet de [JRY00] avec des RTF atomiques et de [NBAMB01] avec des RTF énergétiques. Cette méthode de reconstruction est intéressante pour la séparation de mélanges sous-déterminés pour lesquels il ne suffit pas d'estimer la matrice de mélange pour reconstruire les sources (car \mathbf{A} n'est pas inversible).

3.3 Résultats comparatifs sur des mélanges synthétiques

Nous proposons dans cette partie de comparer les performances de JADE, SOBI, BGML et TFBSS dans plusieurs cas de figure et avec différents réglages des paramètres des méthodes. Auparavant, nous définissons dans le paragraphe suivant les critères d'évaluation utilisés.

3.3.1 Critères d'évaluation

Nous nous intéressons dans ce paragraphe à l'évaluation de la qualité du résultat d'une méthode de séparation, *en simulation*, c-à-d lorsque l'on dispose des sources originales. La mesure de l'erreur entre une source s_k et une estimation \hat{s}_k de cette source est une tâche simple mais non triviale qui n'apparaît pas avoir été très étudiée dans la littérature. Nous proposons une série de critères d'évaluation permettant la mesure de l'erreur totale sur l'estimation d'une source mais aussi la mesure des contributions des différentes causes d'erreur possibles (interférences, bruit, artefacts) [RGF03].

On notera

$$\langle f, g \rangle \stackrel{\text{def}}{=} \frac{1}{T} \sum_{t=0}^{T-1} f[t] \bar{g}[t] \quad (3.58)$$

le produit scalaire entre deux séquences f et g de $\mathbb{C}^{1 \times T}$. La norme de f est alors définie par :

$$\|f\| = \sqrt{\langle f, f \rangle} \quad (3.59)$$

Les séquences f et g sont *orthogonales* ssi $\langle f, g \rangle = 0$.

Il est important de souligner que dans le contexte stochastique qui a été adopté jusqu'à présent, si les sources sont traitées comme étant les réalisations de séquences aléatoires stationnaires ergodiques, le produit scalaire entre deux séquences f et g est égal à l'estimation de la corrélation entre les deux séquences aléatoires dont f et g sont les réalisations. La notion d'*orthogonalité* est alors à rapprocher de la notion de *décorrélation*.

Soit \hat{s}_k une estimation de la source s_k de la forme :

$$\hat{s}_k = \alpha s_k + e \quad (3.60)$$

avec $\alpha \in \mathbb{C}$ et où e désigne un terme d'erreur. La distortion relative entre s_k et \hat{s}_k définie par

$$D_1 \stackrel{\text{def}}{=} \frac{\|\hat{s}_k - s_k\|^2}{\|\hat{s}_k\|^2} \quad (3.61)$$

ne prend pas en compte l'indétermination sur le gain de la SAS. En effet, si \hat{s}_k est une estimation «parfaite» de s_k , donc de la forme $\hat{s}_k = \alpha s_k$, alors la mesure de distortion vaut $D_1 = |1 - \alpha^{-1}|^2$ et elle est donc non nulle.

Dans le cas réel, l'indétermination sur le gain peut être prise en compte en utilisant plutôt la mesure de distortion suivante [STS99] :

$$D_2 \stackrel{\text{def}}{=} \min_{\epsilon=\pm 1} \left\| \frac{\hat{s}_k}{\|\hat{s}_k\|} - \epsilon \frac{s_k}{\|s_k\|} \right\|^2 \quad (3.62)$$

La variable ϵ corrige le signe. Dans le cas complexe, ϵ devrait en plus corriger la phase. Cette mesure garantit bien $D_2 = 0$ si $\hat{s}_k = \alpha s_k$. En revanche, à l'autre extrême, dans le cas où il n'existe aucune contribution d'une source dans son estimation, c-à-d, si $\hat{s}_k = e$ il serait souhaitable que la mesure de distortion soit infinie. Or, si l'erreur e est orthogonale à s_k , la mesure D_2 prend au maximum la valeur 2. On remarquera que la mesure D_2 est «homogène» au produit scalaire des sources, c-à-d, en contexte stochastique, à leur corrélation. En effet, en supposant les sources normalisées et $\epsilon = 1$, on a simplement :

$$D_2 = \|\hat{s}_k\| + \|s_k\| - 2 \operatorname{Re} \left\{ \left\langle \frac{\hat{s}_k}{\|\hat{s}_k\|}, \frac{s_k}{\|s_k\|} \right\rangle \right\} \quad (3.63)$$

$$= 2 (1 - \operatorname{Re} \{ \langle \hat{s}_k, s_k \rangle \}) \quad (3.64)$$

Distortion relative totale

Compte tenu des limitations des distortions D_1 et D_2 , nous proposons de définir la *distortion relative totale* comme le rapport des énergies des deux termes de la décomposition suivante :

$$\hat{s}_k = \left\langle \hat{s}_k, \frac{s_k}{\|s_k\|} \right\rangle \frac{s_k}{\|s_k\|} + e_{total} \quad (3.65)$$

Le terme $\langle \hat{s}_k, s_k / \|s_k\| \rangle s_k / \|s_k\|$ n'est autre que la projection de la source estimée sur la source originale. On définit ainsi la *distortion relative totale* par :

$$D_{total} \stackrel{\text{def}}{=} \frac{\|e_{total}\|^2}{\left\| \left\langle \hat{s}_k, \frac{s_k}{\|s_k\|} \right\rangle \frac{s_k}{\|s_k\|} \right\|^2} = \frac{\|e_{total}\|^2}{\left| \left\langle \hat{s}_k, \frac{s_k}{\|s_k\|} \right\rangle \right|^2} \quad (3.66)$$

Le terme d'erreur e_{total} est par construction orthogonal à la contribution de la source originale s_k . Avec le théorème de Pythagore, on a donc :

$$\|e_{total}\|^2 = \|\hat{s}_k\|^2 - \left\| \left\langle \hat{s}_k, \frac{s_k}{\|s_k\|} \right\rangle \frac{s_k}{\|s_k\|} \right\|^2 \quad (3.67)$$

$$= \|\hat{s}_k\|^2 - \left| \left\langle \hat{s}_k, \frac{s_k}{\|s_k\|} \right\rangle \right|^2 \quad (3.68)$$

D'où :

$$D_{total} = \frac{\|\hat{s}_k\|^2 - \left| \left\langle \hat{s}_k, \frac{s_k}{\|s_k\|} \right\rangle \right|^2}{\left| \left\langle \hat{s}_k, \frac{s_k}{\|s_k\|} \right\rangle \right|^2} \quad (3.69)$$

Cette mesure n'est qu'une légère modification de la mesure D_2 . En effet, dans le cas réel, on a simplement $D_{total} = D_2(4 - D_2)/(2 - D_2)^2$. Si l'on dispose d'une estimation parfaite de s_k ($\hat{s}_k = \alpha s_k$), alors $D_{total} = D_2 = 0$. Cependant, lorsque l'estimation d'une source est orthogonale à la source originale, c-à-d, $\langle \hat{s}_k, s_k \rangle = 0$, alors $D_2 = 2$ tandis que $D_{total} = +\infty$. Nous pensons que cette propriété fait de D_{total} une mesure de distortion plus pertinente que D_2 .

Interférences, Bruit et Artefacts

Le terme d'erreur e_{total} inclut les contributions des autres sources (les interférences), du bruit \mathbf{n} , ainsi que des éventuels «artefacts» dus à l'algorithme de séparation⁴. Dans certains problèmes de SAS, la nature de l'erreur d'estimation est très importante et il convient d'évaluer les différentes contributions des différents types d'erreur. Par exemple, en Séparation de Sources Audio, la distortion due aux artefacts de l'algorithme (provenant par exemple de l'ajout de zéros non-naturels dans la Transformée de Fourier Court-Terme de \hat{s}_k [JRY00]) peut s'avérer être, à l'écoute, plus gênante qu'un résidu de bruit ou d'interférences. Ainsi, nous proposons trois mesures de distortions au lieu d'une seule en décomposant e_{total} en trois termes.

Supposons que les éléments de l'ensemble $\mathcal{E} = \{s_1, \dots, s_n\} \cup \{n_1, \dots, n_m\}$ soient tous mutuellement orthogonaux. Alors l'estimation \hat{s}_k possède la décomposition *orthogonale* suivante :

$$\hat{s}_k = \left\langle \hat{s}_k, \frac{s_k}{\|s_k\|} \right\rangle \frac{s_k}{\|s_k\|} + e_{interf} + e_{bruit} + e_{artef} \quad (3.70)$$

où $\left\langle \hat{s}_k, \frac{s_k}{\|s_k\|} \right\rangle \frac{s_k}{\|s_k\|}$ est la contribution de la vraie source,

$$e_{interf} = \sum_{l \neq k} \left\langle \hat{s}_k, \frac{s_l}{\|s_l\|} \right\rangle \frac{s_l}{\|s_l\|} \quad (3.71)$$

est l'erreur due aux interférences des autres sources,

$$e_{bruit} = \sum_{l=1}^m \left\langle \hat{s}_k, \frac{n_l}{\|n_l\|} \right\rangle \frac{n_l}{\|n_l\|} \quad (3.72)$$

est l'erreur due au bruit additif, et finalement

$$e_{artef} = \hat{s}_k - \left\langle \hat{s}_k, \frac{s_k}{\|s_k\|} \right\rangle \frac{s_k}{\|s_k\|} - e_{interf} - e_{bruit} \quad (3.73)$$

est le résidu éventuel attribué aux artefacts numériques de l'algorithme de séparation.

Dans le cas plus général où les éléments de \mathcal{E} ne sont pas tous mutuellement orthogonaux mais forment cependant un ensemble libre (c-à-d qu'il n'existe pas de combinaison linéaire nulle des sources et/ou des bruits), nous proposons de généraliser la décomposition (3.70) en utilisant différents projecteurs orthogonaux. Soient $P_{\mathbf{s}}$ le projecteur orthogonal sur l'espace engendré par les sources $\{s_1, \dots, s_n\}$, $P_{\mathbf{s}, \mathbf{n}}$ le projecteur orthogonal sur l'espace engendré par les sources et les séquences de bruit \mathcal{E} . La décomposition (3.70) reste valable avec :

$$e_{interf} = P_{\mathbf{s}}(\hat{s}_k) - \left\langle \hat{s}_k, \frac{s_k}{\|s_k\|} \right\rangle \frac{s_k}{\|s_k\|} \quad (3.74)$$

$$e_{bruit} = P_{\mathbf{s}, \mathbf{n}}(\hat{s}_k) - P_{\mathbf{s}}(\hat{s}_k) \quad (3.75)$$

$$e_{artef} = \hat{s}_k - P_{\mathbf{s}, \mathbf{n}}(\hat{s}_k) \quad (3.76)$$

Nous définissons alors la *distortion relative due aux interférences* :

$$D_{interf} = \frac{\|e_{interf}\|^2}{\left\| \left\langle \hat{s}_k, \frac{s_k}{\|s_k\|} \right\rangle \frac{s_k}{\|s_k\|} \right\|^2} = \frac{\|e_{interf}\|^2}{\left| \left\langle \hat{s}_k, \frac{s_k}{\|s_k\|} \right\rangle \right|^2}, \quad (3.77)$$

⁴Les algorithmes fournissant des estimations des sources par application d'une matrice de séparation aux observations ne génèrent pas d'artefacts (voir plus loin).

la *distortion relative due au bruit additif* :

$$D_{bruit} = \frac{\|e_{bruit}\|^2}{\left\| \left\langle \hat{s}_k, \frac{s_k}{\|s_k\|} \right\rangle \frac{s_k}{\|s_k\|} + e_{interf} \right\|^2}, \quad (3.78)$$

et la *distortion relative due aux artefacts* :

$$D_{artef} = \frac{\|e_{artef}\|^2}{\left\| \left\langle \hat{s}_k, \frac{s_k}{\|s_k\|} \right\rangle \frac{s_k}{\|s_k\|} + e_{interf} + e_{bruit} \right\|^2}. \quad (3.79)$$

D'après D_{total} (resp. D_{interf} , D_{bruit} , D_{artef}), on définit en dB le Rapport Source à Distortion (RSD) :

$$\text{RSD} = 10 \log_{10} D_{total}^{-1} \quad (3.80)$$

(resp. Rapport Source à Interférences (RSI), Rapport Signal à Bruit (RSB), Rapport Signal à Artefacts (RSA)).

Une définition plus «naturelle» de la distortion relative due au bruit aurait pu être :

$$D'_{bruit} = \frac{\|e_{bruit}\|^2}{\left\| \left\langle \hat{s}_k, \frac{s_k}{\|s_k\|} \right\rangle \frac{s_k}{\|s_k\|} \right\|^2} = \frac{\|e_{bruit}\|^2}{\left| \left\langle \hat{s}_k, \frac{s_k}{\|s_k\|} \right\rangle \right|^2} \quad (3.81)$$

Cependant telle qu'elle est définie par l'équation (3.78), la mesure D_{bruit} est indépendante de la mesure D_{interf} , contrairement à la mesure D'_{bruit} . Considérons par exemple, un exemple de Séparation de Sources Audio. Soit deux sources s_1 et s_2 orthogonales ($\langle s_1, s_2 \rangle = 0$) et normées ($\|s_1\| = \|s_2\| = 1$). Soit l'estimation suivante de la source s_1 :

$$\hat{s}_1 = \epsilon_1 s_1 + s_2 + e_{bruit} \quad (3.82)$$

où ϵ_1 et $\epsilon_2 = \|e_{bruit}\|$ sont très inférieurs à 1.

Une telle estimation de s_1 est perçue comme pratiquement pas bruitée mais avec beaucoup d'interférences. Ceci est cohérent avec :

$$D_{interf} = \frac{\|s_2\|^2}{\|\epsilon_1 s_1\|^2} = \frac{1}{\epsilon_1^2} \gg 1 \quad (3.83)$$

et

$$D_{bruit} = \frac{\epsilon_2^2}{\|\epsilon_1 s_1 + s_2\|^2} \approx \epsilon_2^2 \ll 1 \quad (3.84)$$

En revanche, la mesure D'_{bruit} vaut :

$$D'_{bruit} = \frac{\epsilon_2^2}{\|\epsilon_1 s_1\|^2} = \frac{\epsilon_2^2}{\epsilon_1^2} \approx 1 \quad (3.85)$$

Avec cette mesure, la contribution du bruit peut donc apparaître comme trop importante, la mesure D_{bruit} nous paraît plus appropriée. De même la définition choisie de la mesure D_{artef} la rend indépendante des mesures D_{bruit} et D_{interf} .

Calcul des mesures de distortion

Le calcul des différentes mesures proposées nécessite le calcul de $\left\langle \hat{s}_k, \frac{s_k}{\|\hat{s}_k\|} \right\rangle$ ainsi que des projections orthogonales $P_{\mathbf{s}}(\hat{s}_k)$ and $P_{\mathbf{s},\mathbf{n}}(\hat{s}_k)$.

Dans le cas de sources mutuellement orthogonales, on a simplement :

$$P_{\mathbf{s}}(\hat{s}_k) = \sum_{l=1}^n \left\langle \hat{s}_k, \frac{s_l}{\|s_l\|} \right\rangle \frac{s_l}{\|s_l\|} \quad (3.86)$$

Dans le cas de sources éventuellement corrélées, le calcul de $P_{\mathbf{s}}(\hat{s}_k)$ est un problème de moindre carrés qui consiste à estimer un vecteur $\mathbf{m}_k \in \mathbb{C}^{n \times 1}$ tel que $P_{\mathbf{s}}(\hat{s}_k) = \mathbf{m}_k^T \mathbf{s}$ et vérifiant :

$$\forall l \in \llbracket 1, n \rrbracket, \langle \hat{s}_k - P_{\mathbf{s}}(\hat{s}_k), s_l \rangle = 0 \quad (3.87)$$

La solution à ce problème fait appel à la matrice de Gram de sources, définie par

$$\mathbf{G}_{\mathbf{s}} = [\langle s_i, s_j \rangle]_{i,j} = \frac{1}{T} \mathbf{s} \mathbf{s}^H \quad (3.88)$$

On a alors :

$$\mathbf{m}_k = \overline{\mathbf{G}_{\mathbf{s}}}^{-1} \mathbf{d}_k \quad (3.89)$$

où $\mathbf{d}_k = [\langle \hat{s}_k, s_1 \rangle, \dots, \langle \hat{s}_k, s_n \rangle]^T$. Encore une fois, si les sources sont les réalisations de séquences aléatoires centrées ergodiques stationnaires, alors la matrice de Gram des sources $\mathbf{G}_{\mathbf{s}}$ correspond à l'estimation de la matrice de covariance $\mathcal{R}_{\mathbf{ss}}[0]$.

La projection $P_{\mathbf{s},\mathbf{n}}(\hat{s}_k)$ peut être calculée de la même façon en considérant la matrice de Gram $\mathbf{G}_{\mathbf{s},\mathbf{n}}$ des sources et des séquences de bruit. Toutefois, si les séquences de bruits sont mutuellement orthogonales et orthogonales aux sources on a plus simplement :

$$P_{\mathbf{s},\mathbf{n}}(\hat{s}_k) = P_{\mathbf{s}}(\hat{s}_k) + \sum_{l=1}^m \left\langle \hat{s}_k, \frac{n_l}{\|n_l\|} \right\rangle \frac{n_l}{\|n_l\|} \quad (3.90)$$

Des routines MATLAB calculant ces mesures de performance sont disponible sur le site Internet de l'Action Jeunes Chercheurs GdR-ISIS «Ressources pour la séparation de sources audio» [GFV].

Un cas particulier

Dans le cas où les estimations des sources proviennent de l'application d'une matrice de séparation \mathbf{B} aux observations \mathbf{x} , et si les sources sont mutuellement orthogonales et orthogonales aux séquences de bruit, le critère D_{interf} peut s'écrire plus simplement en fonction du système global $\mathbf{C} = \mathbf{B} \mathbf{A}$. On a :

$$\hat{\mathbf{s}} = \mathbf{B} \mathbf{x} \quad (3.91)$$

$$= \mathbf{B} \mathbf{A} \mathbf{s} + \mathbf{B} \mathbf{n} \quad (3.92)$$

$$= \mathbf{C} \mathbf{s} + \mathbf{B} \mathbf{n} \quad (3.93)$$

On suppose que l'indétermination sur l'ordre des sources a été levée ; la séparation est donc parfaite si \mathbf{C} est diagonale. Avec $\mathbf{C} = \{c_{ij}\}$ et $\mathbf{B} = \{b_{ij}\}$, on a :

$$\hat{s}_k = \sum_{l=1}^n c_{kl} s_l + \sum_{l=1}^m b_{kl} n_l \quad (3.94)$$

D'où :

$$e_{interf} = \sum_{l \neq k} c_{kl} s_l \quad (3.95)$$

$$e_{bruit} = \sum_{l=1}^m b_{kl} n_l \quad (3.96)$$

$$e_{artef} = \mathbf{0} \quad (3.97)$$

D'où :

$$\|e_{interf}\|^2 = \sum_{l \neq k} |c_{kl}|^2 \|s_l\|^2 \quad (3.98)$$

et

$$\left\langle \hat{s}_k, \frac{s_k}{\|s_k\|} \right\rangle = \left\langle c_{kk} s_k, \frac{s_k}{\|s_k\|} \right\rangle \quad (3.99)$$

$$= c_{kk} \|s_k\| \quad (3.100)$$

D'après l'équation (3.77), on a donc :

$$D_{interf} = \frac{\sum_{l \neq k} |c_{kl}|^2 \|s_l\|^2}{|c_{kk}|^2 \|s_k\|^2} \quad (3.101)$$

D_{interf} s'exprime donc simplement en fonction des termes de la k^{eme} ligne de \mathbf{C} .

Pour donner une idée de ce que représente une valeur donnée en dB du RSI, fixons l'indétermination sur le gain en supposant les sources normées : $\forall i, \|s_i\| = 1$. Si la séparation est parfaite, la matrice \mathbf{C} est donc égale à l'identité. On peut écrire :

$$\sum_{l \neq k} |c_{kl}|^2 = 10^{-\frac{RSI}{10}} |c_{kk}|^2 \quad (3.102)$$

Avec $\|\cdot\|_1 \leq \sqrt{\dim(\cdot)} \|\cdot\|_2$, il vient :

$$\sum_{l \neq k} |c_{kl}| \leq \sqrt{n-1} 10^{-\frac{RSI}{20}} |c_{kk}| \quad (3.103)$$

Par exemple, si $RSI = 30$ dB et $n = 3$, alors :

$$\sum_{l \neq k} |c_{kl}| \leq 0.0447 |c_{kk}| \quad (3.104)$$

D'un point de vue auditif, au-delà d'un RSI de 30 dB, il est impossible de distinguer les interférences dans l'estimation d'une source audio.

Enfin, notons que les hypothèses formulées dans ce paragraphe sont le pendant idéal en pratique des hypothèses stochastiques formulées sur les sources et sur le bruit dans les méthodes exposées aux précédents chapitres (sources mutuellement indépendantes au moins à l'ordre 2 et indépendantes du bruit). Si les sources ne sont pas rigoureusement orthogonales ni les séquences bruit rigoureusement orthogonales aux sources, on peut néanmoins s'attendre à une valeur de D_{interf} proche de celle donnée par (3.101).

Conclusion

Nous avons présenté dans ce paragraphe plusieurs mesures d'erreur, une mesure d'erreur totale D_{total} qui permet d'évaluer la qualité générale de la séparation mais aussi trois mesures D_{interf} , D_{bruit} et D_{artef} permettant d'évaluer les contributions des différentes sources possibles d'erreur, à savoir les interférences, le bruit et les artefacts. L'intérêt de ces critères est mise en évidence sur des résultats en simulation dans le paragraphe suivant sur des mélanges linéaires instantanés ainsi qu'au paragraphe 4.3 sur des mélanges convolutifs.

3.3.2 Performances comparées

Nous présentons dans ce paragraphe une série de résultats sur des mélanges synthétiques. Dans un premier temps nous proposons une étude statistique des performances comparées de TFBSS, SOBI et JADE sur des signaux synthétiques non-stationnaires TVARMA, en fonction du niveau de bruit sur les observations et du recouvrement spectral des sources⁵. Puis, nous étudierons les performances de la méthode temps-fréquence selon le critère de sélection des points t-f (sélection d'auto-termes simples ou des points de plus haute énergie), pour différents paramètres de réglages. L'influence du choix de la RTFS sera notamment étudiée. Les performances de TFBSS et BGML seront aussi comparées, en contexte déterminé ($m = n$) et non-bruité. Nous discuterons aussi de l'influence de la longueur des signaux selon leur type (stationnaires ou non-stationnaires) sur les performances de TFBSS, SOBI et JADE.

Enfin, nous présenterons et discuterons les résultats de séparation d'un mélange synthétique de sources audio.

Définition du Rapport Signal à Bruit d'Entrée

Les performances de TFBSS, SOBI et JADE sont étudiées en contexte bruité. Nous définissons ici le Rapport Signal à Bruit d'Entrée (RSBE) utilisé par la suite. Une observation quelconque x_i ($i \in \llbracket 1, m \rrbracket$) s'écrit :

$$x_i = \sum_{j=1}^n a_{ij} s_j + n_i \quad (3.105)$$

Nous définissons le RSBE de l'observation x_i par :

$$\text{RSBE}_{x_i} \stackrel{\text{def}}{=} 10 \log \frac{\sum_{j=1}^n |a_{ij}|^2 \|s_j\|^2}{\|n_i\|^2} \quad (3.106)$$

En supposant les sources normées ($\|s_i\| = 1$) et avec $\|n_i\| \approx \sigma$, on a plus simplement :

$$\text{RSBE}_{x_i} \approx 20 \log \frac{\sqrt{\sum_{j=1}^n |a_{ij}|^2}}{\sigma} \quad (3.107)$$

Le RSBE général est défini par :

$$\text{RSBE} \stackrel{\text{def}}{=} \min_i \{\text{RSBE}_{x_i}\} \quad (3.108)$$

Comparaison de TFBSS, SOBI et JADE

Nous comparons les performances de TFBSS, SOBI et JADE sur des mélanges de sources de type TVARMA (Time-Varying ARMA). Ce sont des séquences i.i.d filtrées par un filtre ARMA dont les coefficients varient au cours du temps. Les sources sont donc les réalisations de séquences non-stationnaires.

⁵On entendra par «recouvrement spectral» des sources le degré de recouvrement des signatures t-f des sources.

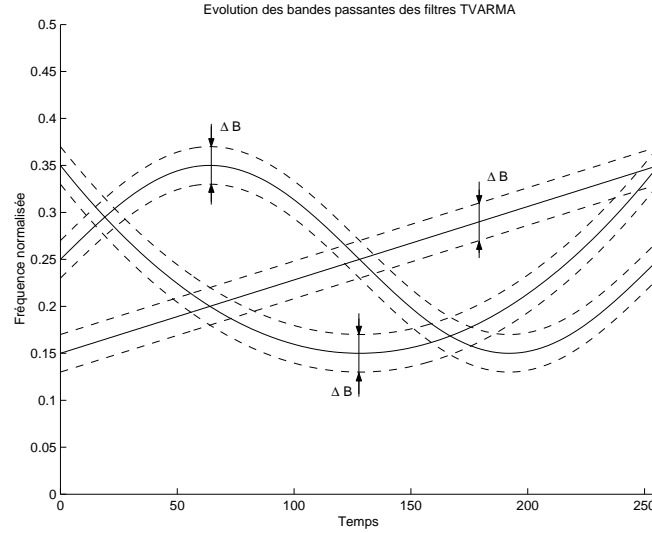


FIG. 3.1 – Les sources sont générées par des séquences i.i.d gaussiennes filtrées par des filtres ARMA dont les fréquences de coupure varient dans le temps selon différents profils géométriques (droite, sinus, parabole). Le paramètre ΔB règle la bande passante.

Les séquences d'entrée des filtres sont choisies gaussiennes de longueur $T = 256$. Les filtres TVARMA sont des filtres de Butterworth passe-bande d'ordre 8 dont les fréquences de coupure varient au cours du temps. On génère $n = 3$ sources suivant 3 profils d'évolution des fréquences de coupure, représentés sur la figure 3.1. Le paramètre ΔB règle la bande passante des filtres et détermine donc le recouvrement spectral des sources.

Les figures 3.2 et 3.3 présentent deux exemples de réalisations des sources normalisées pour $\Delta B = 0.08$ et $\Delta B = 0.24$, accompagnées d'une RTF (distribution pseudo-Wigner-Ville lissée, voir plus loin).

Pour 3 cas de recouvrement spectral ($\Delta B \in \{0.08, 0.16, 0.24\}$) et différents niveaux de bruit ($\text{RSBE} \in \{0, 5, 10, 15, 20, 25, 30, 35, 40\}$), nous avons généré $N_{sim} = 500$ réalisations des sources et nous avons créé pour chaque réalisation un mélange synthétique comportant $m = 4$ observations. La matrice de mélange est fixée et vaut :

$$\mathbf{A} = \begin{bmatrix} 1 & 0.5 & 0.3 \\ 0.6 & 1 & 0.5 \\ 0.4 & 0.8 & 1 \\ 0.4 & 0.6 & 0.3 \end{bmatrix} \quad (3.109)$$

TFBSS, SOBI et JADE ont été appliqués à chaque réalisation des observations et nous avons calculé les valeurs moyennes des RSD, RSI et RSB des estimations des trois sources. On notera :

$$\text{RSD} = \frac{1}{n} \sum_{i=1}^n \text{RSD}_{\hat{s}_i} \quad (3.110)$$

où $\text{RSD}_{\hat{s}_i}$ désigne le RSD de l'estimation \hat{s}_i de la source s_i (on définit de même RSI et RSB). Enfin, pour chaque valeur de ΔB et RSBE nous calculons la moyenne et l'écart type de RSD , RSI et RSB sur les 500 réalisations.

La RTFS des observations utilisée dans TFBSS est une distribution pseudo-Wigner-Ville lissée, calculée sur $I_T \times I_F = \llbracket 0, T-1 \rrbracket \times \llbracket 0, T-1 \rrbracket$. La sélection des auto-termes simples des sources a été réalisée avec $\varepsilon_{Grad} = 0.01$. L'influence de ces différents paramètres (choix de la distribution, nombre de points dans I_T et I_F , valeur de ε_{Grad}) est étudiée par la suite. L'ensemble des retards utilisés dans SOBI est fixé à $\{1, 2, 3\}$.

Les résultats sont reportés sur les figures 3.4, 3.5 et 3.6.

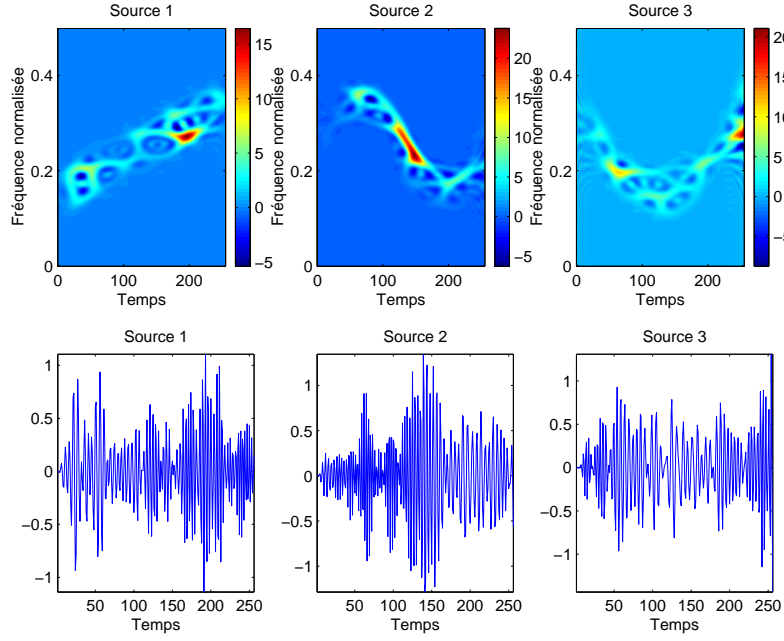


FIG. 3.2 – Réalisations temporelles de chaque source pour $\Delta B = 0.08$, accompagnées de leur distribution pseudo-Wigner-Ville lissée.

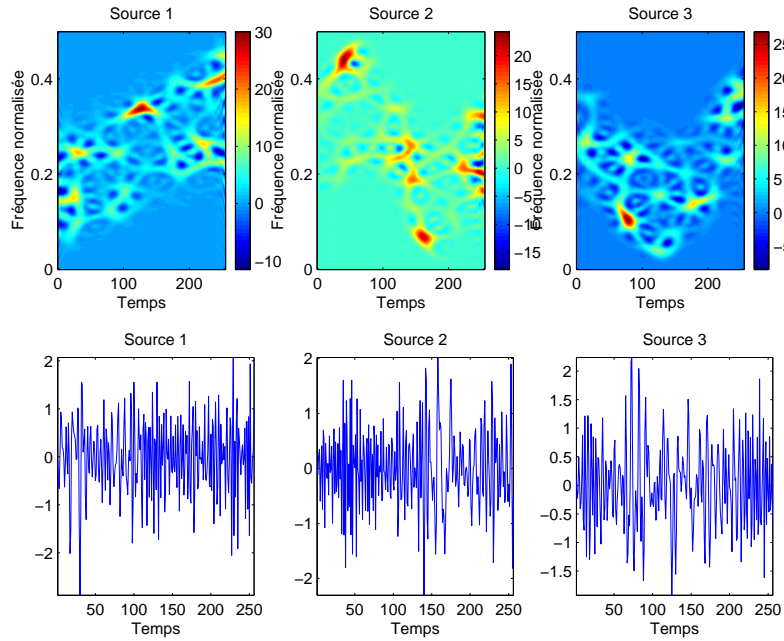


FIG. 3.3 – Réalisations temporelles de chaque source pour $\Delta B = 0.24$, accompagnées de leur distribution pseudo-Wigner-Ville lissée.

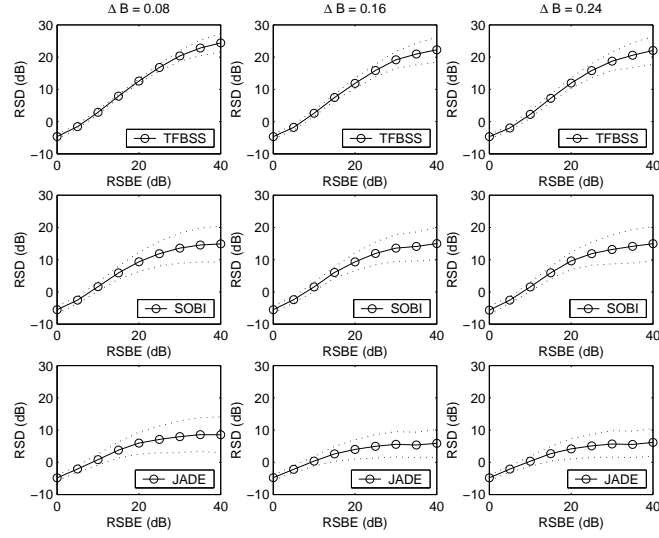


FIG. 3.4 – Comparaison de TFBSS, JADE et SOBI sur des mélanges de sources TVARMA pour différents cas de recouvrement spectral (en colonnes) et différents niveaux de bruit d'entrée. Valeur moyenne ($\circ-$) \pm écart type ($\cdot\cdot$) de \bar{RSD} sur 500 réalisations.

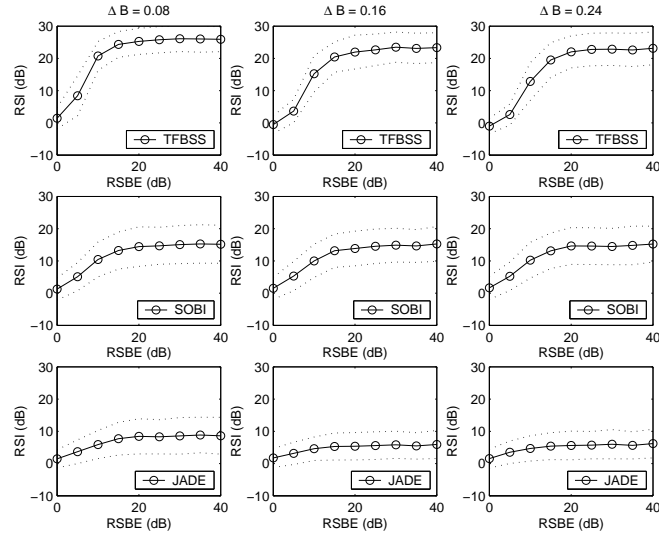


FIG. 3.5 – Comparaison de TFBSS, JADE et SOBI sur des mélanges de sources TVARMA pour différents cas de recouvrement spectral (en colonnes) et différents niveaux de bruit d'entrée. Valeur moyenne ($\circ-$) \pm écart type ($\cdot\cdot$) de \bar{RSI} sur 500 réalisations.

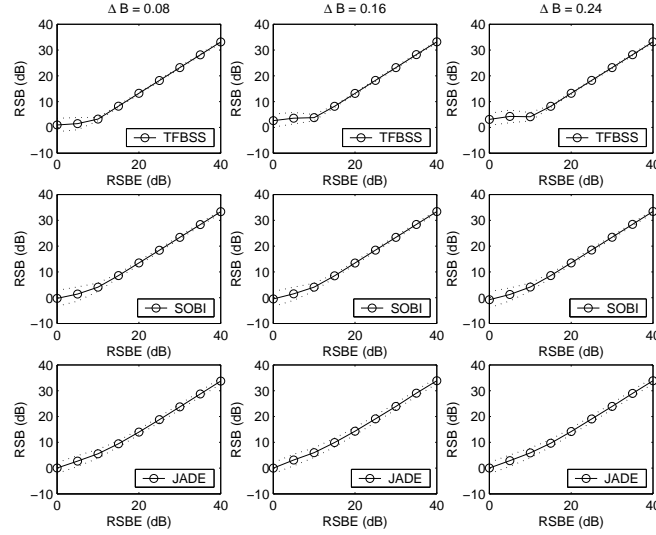


FIG. 3.6 – Comparaison de TFBSS, JADE et SOBI sur des mélanges de sources TVARMA pour différents cas de recouvrement spectral (en colonnes) et différents niveaux de bruit d'entrée. Valeur moyenne (\circ) \pm écart type ($\cdot\cdot$) de $\bar{R}SB$ sur 500 réalisations.

Discussion La figure 3.6 montre que la contribution du bruit dans l'erreur d'estimation des sources ($\bar{R}SB$) varie linéairement avec le niveau de bruit sur les observations (RSBE). Ce qui est cohérent avec le fait que du bruit additif subsiste sur les estimations des sources (équation (2.3)). Il est alors logique que l'erreur totale sur l'estimation des sources ($\bar{R}SD$) chute sensiblement lorsque le niveau de bruit en entrée augmente (c-à-d, $RSBE \rightarrow 0$), comme le montre la figure 3.4. Il est donc plus pertinent d'étudier les variations du RSI (figure 3.5), qui mesure uniquement le pouvoir de réjection des autres sources dans l'estimation d'une source.

On constate alors que les 3 méthodes sont, en termes de réjection, robustes au bruit sur l'intervalle $RSBE = [10, 40]$, puis leurs performances chutent sensiblement.

Les performances des trois méthodes sont indifférentes au recouvrement spectral. TFBSS fournit les meilleures performances sur la plage $RSBE = [10, 40]$, avec un gain de 10 dB par rapport aux performances de SOBI. Notons que SOBI donne de bons résultats, bien que cet algorithme soit construit pour des sources stationnaires. En revanche, les résultats de JADE sont décevants.

Comparaison des méthodes de sélection des points t-f

Nous étudions dans ce paragraphe les performances de la méthode t-f suivant le mode de sélection des points : sélection d'Auto-Terms Simples des sources ou sélection des Points de plus Haute Énergie. La première méthode de sélection est nommée «méthode ATS», la deuxième «méthode PHE».

Dans un premier temps, une étude statistique des performances des deux modes de sélection est menée sur le même type de sources et la même matrice de mélange que précédemment, avec $N_{sim} = 100$ réalisations.

Nous avons étudié l'influence du nombre de points temporels (N_T) et fréquentiels (N_F) utilisés pour le calcul de la RTFS des observations. Ces points sont choisis de manière équirépartie

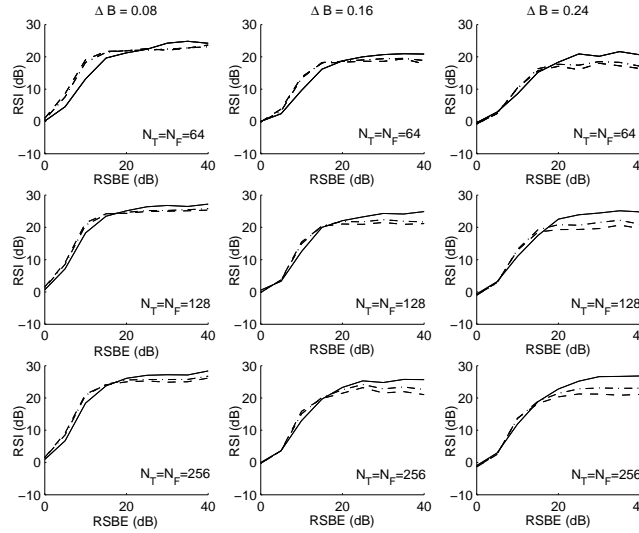


FIG. 3.7 – Résultats de la méthode t-f en utilisant la méthode de sélection ATS, pour 3 cas de recouvrement spectral (en colonnes). $(-)$ $\varepsilon_{Grad} = \frac{2}{N_T + N_F}$, $(- -)$ $\varepsilon_{Grad} = \frac{10}{N_T + N_F}$, $(- \cdot -)$ $\varepsilon_{Grad} = \frac{20}{N_T + N_F}$.

sur les intervalles $\llbracket 0, T - 1 \rrbracket \times \llbracket 0, T - 1 \rrbracket$, c-à-d (en supposant T/N_T et T/N_F entiers) :

$$I_T = \left\{0, \frac{T}{N_T}, 2 \frac{T}{N_T}, 3 \frac{T}{N_T}, \dots\right\} \quad (3.111)$$

$$I_F = \left\{0, \frac{T}{N_F}, 2 \frac{T}{N_F}, 3 \frac{T}{N_F}, \dots\right\} \quad (3.112)$$

Par exemple, avec $T = 256$, si $N_T = N_F = 64$, on calcule $\hat{\mathcal{S}}_{xx}^\phi[t, f]$ pour $(t, f) \in \{0, 4, 8, \dots, 252\}^2$. Nous avons étudié trois cas : $(N_T = N_F) \in \{64, 128, 256\}$.

Pour chaque valeur de (N_T, N_F) nous avons appliqué les méthodes aux signaux en testant plusieurs valeurs des seuils ε_{Grad} et ε_{Energ} des méthodes de sélection. Le seuil ε_{Grad} étant homogène au gradient de $C_2[t, f]$, nous l'avons choisi proportionnellement à $1/(N_T + N_F)$. La figure 3.7 présente les résultats obtenus avec la méthode ATS pour $\varepsilon_{Grad} \in \{2, 10, 20\}/(N_T + N_F)$ ⁶. La figure 3.8 présente les résultats obtenus avec la méthode PHE pour $\varepsilon_{Energ} \in \{0.9, 1, 1.1\}$. La figure 3.9 compare les résultats obtenus par les deux méthodes de sélection pour les choix optimaux (parmi les valeurs testées) $\varepsilon_{Grad} = 2/(N_T + N_F)$ et $\varepsilon_{Energ} = 0.9$. La figure 3.10 trace dans ce dernier cas le nombre moyen de points t-f sélectionnés par chacune des méthodes de sélection.

Dans un deuxième temps, nous proposons de voir sur un exemple simple où sont localisés dans le plan t-f les points sélectionnés par les deux méthodes. A cette fin nous considérons $n = 3$ sources déterministes à bande étroite : 3 sinusoides modulées en fréquence suivant un profil linéaire, un profil sinusoidal et un profil parabolique. Nous avons généré $m = 4$ observations avec \mathbf{A} définie par l'équation (3.109) et RSBE=20 dB.

La figure 3.11 (a) représente un critère de diagonalité des matrices $\hat{\mathcal{S}}_{ss}^\phi[t, f]$ pour $(t, f) \in \llbracket 0, T - 1 \rrbracket^2$. Le critère de diagonalité utilisé est :

$$C_d(\mathbf{M}) = \frac{\sum_i |m_{ii}|}{\sum_{ij} |m_{ij}|} \quad (3.113)$$

⁶En pratique nous avons constaté que des valeurs de ε_{Grad} inférieures à $2/(N_T + N_F)$ peuvent poser problème : en raison de l'implémentation discrète du gradient, si le seuil est trop faible il est possible qu'aucun point t-f ne soit sélectionné.

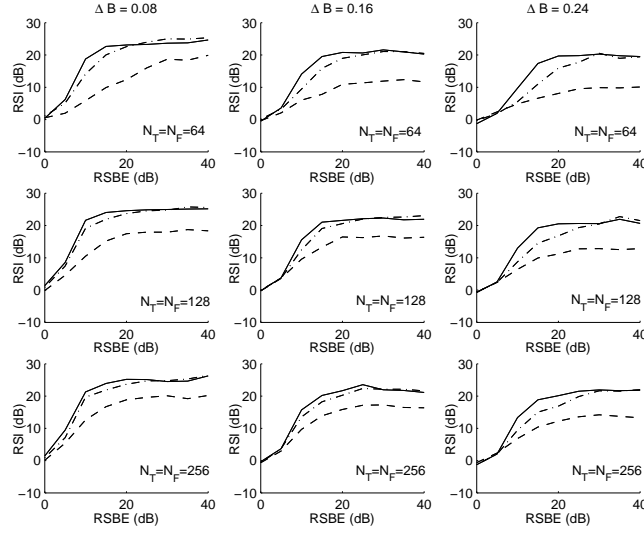


FIG. 3.8 – Résultats de la méthode t-f en utilisant la méthode de sélection PHE, pour 3 cas de recouvrement spectral (en colonnes). (—) $\varepsilon_{Energ} = 0.9$, (---) $\varepsilon_{Energ} = 1$, (-.-) $\varepsilon_{Energ} = 1.1$.

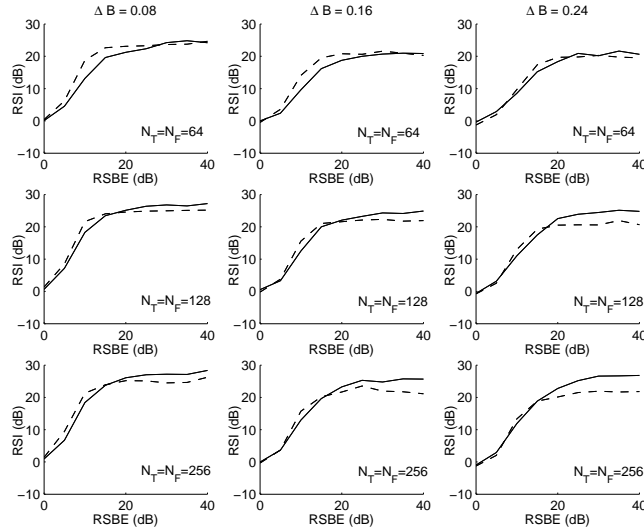


FIG. 3.9 – Résultats comparés des méthodes de sélection pour les choix optimaux de ε_{Grad} et ε_{Energ} , pour 3 cas de recouvrement spectral (en colonnes). (—) Méthode ATS avec $\varepsilon_{Grad} = \frac{2}{N_T + N_F}$, (---) Méthode PHE avec $\varepsilon_{Energ} = 0.9$.

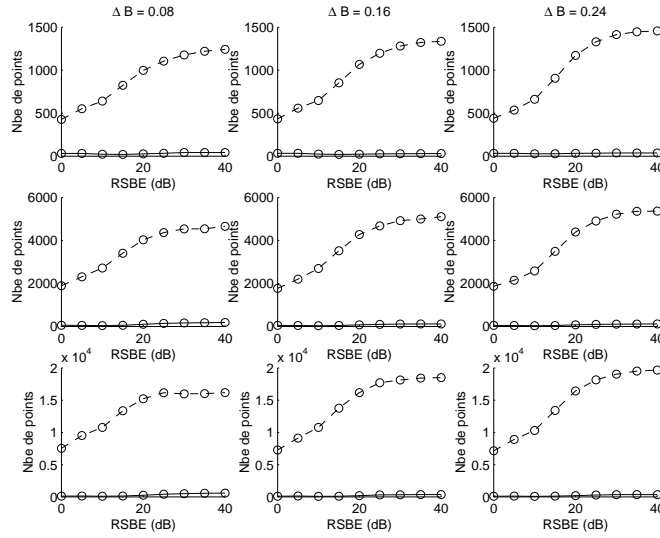


FIG. 3.10 – Nombre moyen de points t-f sélectionnés par chacune des deux méthodes pour les choix optimaux de ε_{Grad} et ε_{Energ} . (—) Méthode ATS avec $\varepsilon_{Grad} = \frac{2}{N_T + N_F}$, (---) Méthode PHE avec $\varepsilon_{Energ} = 0.9$.

où $\mathbf{M} = \{m_{ij}\} \in \mathbb{C}^{n \times n}$. Afin de vérifier que les matrices $\hat{\mathcal{S}}_{ss}^\phi[t, f]$ diagonales n'ont qu'un seul terme diagonal dominant, la figure 3.11 (b) représente le critère suivant :

$$C_{1d}(\mathbf{M}) = \frac{\max_i |m_{ii}|}{\sum_{ij} |m_{ij}|} \quad (3.114)$$

$C_{1d}(\hat{\mathcal{S}}_{ss}^\phi[t, f])$ est proche de 1 si $\hat{\mathcal{S}}_{ss}^\phi[t, f]$ est diagonale avec un seul terme dominant.

Les zones blanches de la figure 3.11 (a) correspondent à $C_d(\hat{\mathcal{S}}_{ss}^\phi[t, f]) \approx 1$, c-à-d, aux matrices $\hat{\mathcal{S}}_{ss}^\phi[t, f]$ les plus diagonales. Le but des méthodes de sélection est donc de sélectionner des points appartenant à ces zones blanches. Les figures 3.11 (a) et (b) montrent que les maxima de $C_d(\hat{\mathcal{S}}_{ss}^\phi[t, f])$ sont égaux aux maxima de $C_{1d}(\hat{\mathcal{S}}_{ss}^\phi[t, f])$, ce qui montre que les matrices $\hat{\mathcal{S}}_{ss}^\phi[t, f]$ diagonales n'ont bien qu'un seul terme dominant sur la diagonale et souligne la pertinence de rechercher des auto-termes simples des sources.

La figure 3.11 (c) représente les points sélectionnés par la méthode de sélection des auto-termes simples, avec $\varepsilon_{Grad} = 2/(N_T + N_F)$, la figure 3.11 (d) représente les points sélectionnés par la méthode de sélection des points de plus haute énergie, avec $\varepsilon_{Energ} = 0.9$.

Discussion Les résultats fournis par la méthode de sélection ATS (figure 3.7) sont peu sensibles au choix du seuil ε_{Grad} , ni au nombre de points t-f $N_T \times N_F$ et au recouvrement spectral. En revanche, la méthode de sélection PHE est très sensible au choix du seuil ε_{Energ} . Les performances de la méthode chutent sensiblement de $\varepsilon_{Energ} = 0.9$ à $\varepsilon_{Energ} = 1.1$ (voir figure 3.8). Nous proposons une explication à ce phénomène. Idéalement, les auto-termes simples des sources vérifient :

$$\text{trace}(\hat{\mathcal{S}}_{ss}^\phi[t, f]) = \left\| \hat{\mathcal{S}}_{ss}^\phi[t, f] \right\|_F \quad (3.115)$$

Donc, d'après l'expression (3.47) du critère de sélection PHE, si $\varepsilon_{Energ} > 1$, les points t-f associés à des auto-termes simples des sources ne sont pas sélectionnés par la méthode PHE et ceci expliquerait les mauvaises performances de cette méthode pour $\varepsilon_{Energ} = 1.1$.

La comparaison des méthodes de sélection pour les choix optimaux de ε_{Grad} et ε_{Energ} révèle des résultats similaires. Cependant, le nombre moyen de points t-f sélectionnés par chacune des méthodes est incomparable : il varie de 30 à 500 pour la méthode ATS, alors qu'il varie de 500 à

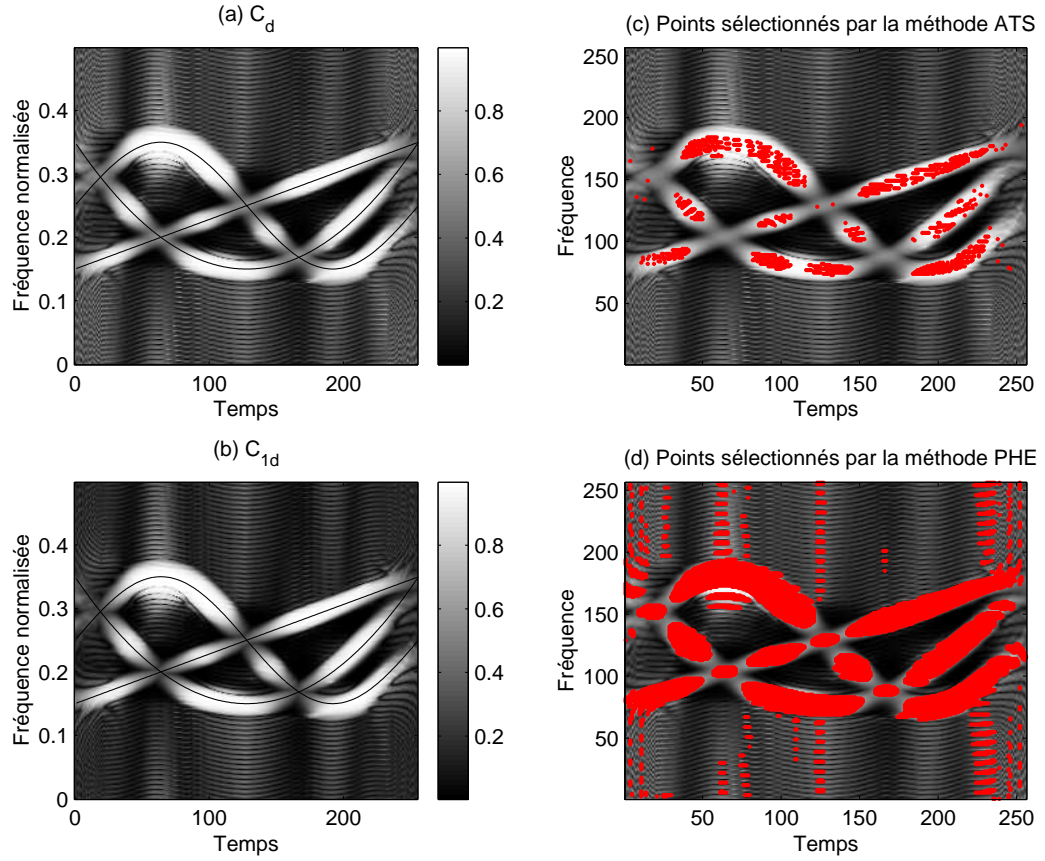


FIG. 3.11 – Localisation des points t-f sélectionnés par deux méthodes. (a) Critère de diagonalité $C_d(\hat{S}_{ss}^\phi[t, f])$, les zones blanches correspondent aux points où $\hat{S}_{ss}^\phi[t, f]$ est la plus diagonale; les traits pleins représentent les fréquences instantanées des sources. (b) Critère de diagonalité avec un terme diagonal dominant $C_{1d}(\hat{S}_{ss}^\phi[t, f])$, (c) Points sélectionnés par la méthode ATS (en rouge), superposés à $C_d(\hat{S}_{ss}^\phi[t, f])$, (d) Points sélectionnés par la méthode PHE (en rouge), superposés à $C_d(\hat{S}_{ss}^\phi[t, f])$.

20000 pour la méthode PHE. Ceci se traduit par une différence de temps de calcul considérable entre les deux méthodes. Dans le cas $N_T = N_F = 256$, sur un ordinateur Pentium 4 cadencé à 2.8 GHz, le temps de calcul moyen de la méthode de séparation t-f globale est de 8s avec la méthode ATS et de 60s avec la méthode PHE.

La figure 3.11 (d) montre que la méthode PHE sélectionne de nombreux point t-f correspondant à des matrices $\hat{S}_{ss}^\phi[t, f]$ diagonales mais aussi de nombreux points «de nuisance». En particulier, la plupart des points localisés à l'intersection des signatures t-f de deux sources sont sélectionnés par la méthode PHE alors qu'il ne correspondent pas à des matrices $\hat{S}_{ss}^\phi[t, f]$ diagonales. L'influence de ces points de nuisance est réduite par la diagonalisation simultanée, dans la mesure où la majorité des points t-f sélectionnés correspondent à des matrices $\hat{S}_{ss}^\phi[t, f]$ effectivement diagonales. Ceci explique que la méthode PHE donne néanmoins de bons résultats. En revanche, la figure 3.11 (c) montre que la méthode ATS ne sélectionne que peu de points t-f, qui correspondent bien tous à des matrices $\hat{S}_{ss}^\phi[t, f]$ diagonales.

Etude de l'influence du noyau

Nous étudions désormais l'influence de la RTFS choisie dans TFBSS. Nous avons repris le protocole de simulation précédent en testant plusieurs RTFS de la classe de Cohen, calculées au moyen de la Boîte à Outils Temps-Fréquence pour MATLAB [AFLG] et utilisées avec leur paramètres par défaut :

- distribution de Wigner-Ville (`tfwv.m`) : $\phi[t, f] = \delta[t] \delta[f]$ (pas de lissage),
- distribution pseudo-Wigner-Ville lissée (`tferspwm.m`) : $\phi[t, f] = h[t] g[f]$ (lissages temps et fréquence disjoints),
- distribution à Interférences Réduites avec noyau de Bessel (`tfrridb.m`), lissage conjoint en temps et en fréquence,
- distribution à Interférences Réduites avec noyau de Choi-Williams (`tfrcw.m`), lissage conjoint en temps et en fréquence.

Une description complète des noyaux utilisés est disponible dans [Fla99, AFLG96].

Les résultats sont présentés sur la figure 3.12 et sont basés sur $N_{sim} = 100$ réalisations des sources pour chaque cas de recouvrement spectral et de RSBE. La méthode de sélection des points t-f est la méthode ATS avec $\varepsilon_{Grad} = \frac{10}{N_T + N_F}$ et $N_T = N_F = 128$.

Discussion Les courbes de la figure 3.12 montrent que les performances de la méthode t-f sont similaires pour les trois distributions SPWV, RIDB et CW. Les performances de la méthode avec la distribution de Wigner-Ville sont légèrement moins bonnes, en particulier la méthode devient moins robuste au bruit. On retiendra donc que les performances de TFBSS sont similaires selon le noyau utilisé du moment que la distribution de Wigner-Ville (spatiale) est lissée en temps et en fréquence. La distribution pseudo-Wigner-Ville lissée étant la moins coûteuse en calculs, c'est celle qui est utilisée en pratique.

Comparaison de l'approche temps-fréquence et BGML

Dans ce paragraphe nous souhaitons mettre en valeur l'apport de l'approche t-f sur la méthode BGML. Les deux méthodes sont comparées sur le même type de sources que précédemment mais cette fois de longueur⁷ $T = 2048$. Nous ne considérons ici que le cas déterminé ($m = n = 3$) non-bruité. La matrice de mélange est choisie arbitrairement et vaut :

$$\mathbf{A} = \begin{bmatrix} 1 & 0.5 & 0.3 \\ 0.6 & 1 & 0.5 \\ 0.4 & 0.8 & 1 \end{bmatrix} \quad (3.116)$$

⁷Ceci afin de pouvoir tester plusieurs longueurs de fenêtre, voir plus loin.

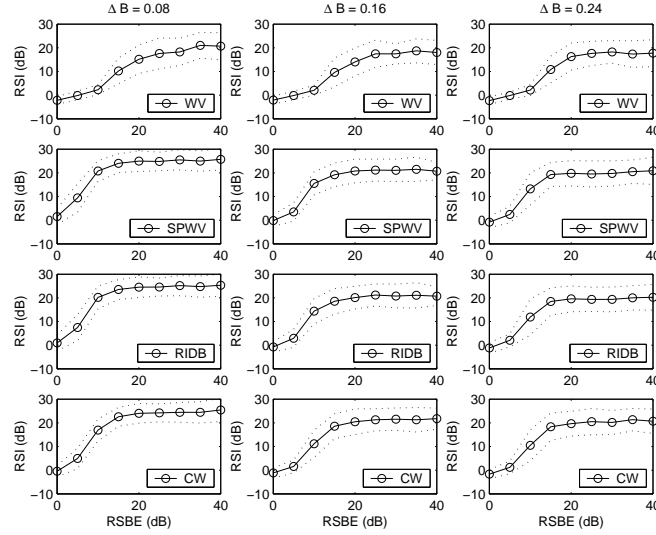


FIG. 3.12 – Comparaison des performances de TFBSS pour différentes RTFS : distribution de Wigner-Ville (WV), distribution pseudo-Wigner-Ville lissée (SPWV), distribution à Interférences Réduites avec noyau de Bessel (RIDB), distribution à Interférences Réduites avec noyau de Choi-Williams (CW), pour 3 cas de recouvrement spectral (en colonnes). Valeur moyenne et écart-type de RSI sur 100 réalisations.

Nous avons étudié les performances de BGML suivant le nombre K de segments en lesquels les observations sont découpées (voir paragraphe 3.1). La longueur des intervalles ayant été choisie comme constante, ceci revient à choisir différentes valeurs de la longueur T_{fen} de la fenêtre (rectangulaire) utilisée pour le calcul des covariances locales. On a donc $K = \left\lfloor \frac{T}{T_{fen}} \right\rfloor$, où $\lfloor \cdot \rfloor$ désigne la partie entière. Nous avons testé les valeurs $T_{fen} \in \{17, 33, 65, 129, 257\}$ (le choix de valeurs impaires permet l'existence d'un instant central, utilisé par la suite).

Pour chaque valeur de T_{fen} nous avons comparé BGML à une version particulière de TFBSS, notée TFBSS'. En effet, la distribution t-f utilisée est la distribution pseudo-Wigner-Ville lissée, mais avec une fenêtre de lissage temporel identique à celle de BGML, c-à-d, une fenêtre rectangulaire de longueur T_{fen} . La RTFS des observations est calculée sur les K instants centraux des K intervalles utilisées par BGML, de sorte que le «découpage temporel» des observations est le même pour les deux méthodes. Le nombre de points fréquentiels pour le calcul de la RTFS des observations est fixé à $N_F = 256$, quelle que soit la valeur de T_{fen} .

Pour chaque valeur de T_{fen} , les algorithmes de séparation ont été appliqués à $N_{sim} = 100$ réalisations des observations. La figure 3.13 présente la valeur moyenne et l'écart-type de RSI sur les 100 réalisations. Pour comparaison, nous présentons aussi les résultats fournis par SOBI pour plusieurs valeurs du nombre de retards.

Discussion La figure 3.13 montre que le gain de performance apporté par l'approche t-f est considérable. On peut constater dans un premier temps que TFBSS' est peu sensible à la longueur de la fenêtre temporelle, à la différence de BGML. Les meilleurs résultats de BGML sont fournis dans les 3 cas de recouvrement spectral pour la longueur de fenêtre la plus courte, c-à-d, $T_{fen} = 17$.

Pour cette longueur de fenêtre, les performances de BGML sont relativement proches de celles de TFBSS' dans le cas $\Delta B = 0.08$ (faible recouvrement spectral). En revanche, lorsque le recouvrement spectral augmente, les performances de BGML se dégradent alors que celles de TFBSS' restent stables. Dans le plus fort cas de recouvrement spectral, les performances de SOBI sont même légèrement meilleures à celles de BGML.

Enfin on remarquera que les performances de SOBI sont peu sensibles au nombre de retards

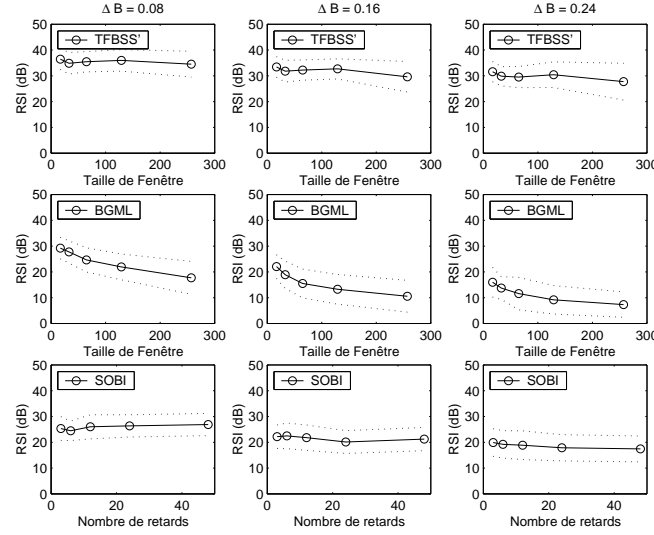


FIG. 3.13 – Performances de TFBSS' et BGML pour plusieurs longueurs de T_{fen} et performance de SOBI pour différents nombres de retards, pour 3 cas de recouvrement spectral (en colonnes). Valeur moyenne et écart-type de RSI sur 100 réalisations.

choisis, c-à-d, au nombre de matrices de covariances diagonalisées simultanément.

En conclusion, ces résultats dégagent l'intérêt de l'approche t-f, c-à-d, de l'utilisation de l'information temps *et* fréquence contenue dans les observations (TFBSS') par rapport à la seule information temporelle (BGML). L'approche t-f est toutefois évidemment beaucoup plus coûteuse en temps de calcul : sur un ordinateur Pentium 4 cadencé à 2.8 GHz, le temps de calcul de la méthode BGML sur des signaux de longueur $T = 2048$ est de l'ordre de 1/10s avec $T_{fen} = 17$, alors qu'il est de l'ordre de 9s pour TFBSS avec $N_T = N_F = 256$ et $\varepsilon_{Grad} = 0.01$.

Etude de l'influence de la longueur des signaux

Nous nous intéressons dans ce paragraphe à l'influence de la longueur des signaux sur les performances de TFBSS, SOBI et JADE. Deux cas de figure sont étudiés :

- 1) Dans un premier temps nous considérons des sources non-stationnaires. Afin de conserver le même «type» de non-stationnarité au cours du temps, on considère $n = 2$ sources TVARMA dont le contenu fréquentiel est périodique dans le temps. Les sources sont encore générées par des séquences gaussiennes filtrées par des filtres de Butterworth d'ordre 8 dont les fréquences de coupures évoluent dans le temps. Une période de l'évolution des fréquences de coupures de chaque source est représentée sur la figure 3.14. Cette période est de 256 points temporels.

Les 3 méthodes sont testées sur $N_{sim} = 100$ réalisations des sources, pour 3 cas de recouvrement spectral ($\Delta B \in \{0.08, 0.16, 0.24\}$) et plusieurs valeurs du nombre de périodes d'évolution fréquentielle des sources, équivalentes en nombres de points à $T \in \{256, 1024, 4096, 16384\}$.

- 2) Dans un second temps nous considérons des sources stationnaires, générées par des séquences i.i.d de densité uniforme filtrées par des filtres de Butterworth passe-bas d'ordre 8 et de fréquences de coupure normalisées 0.3 et 0.45. Les 3 méthodes sont testées sur $N_{sim} = 100$ réalisations des sources et pour $T \in \{256, 1024, 4096, 16384\}$.

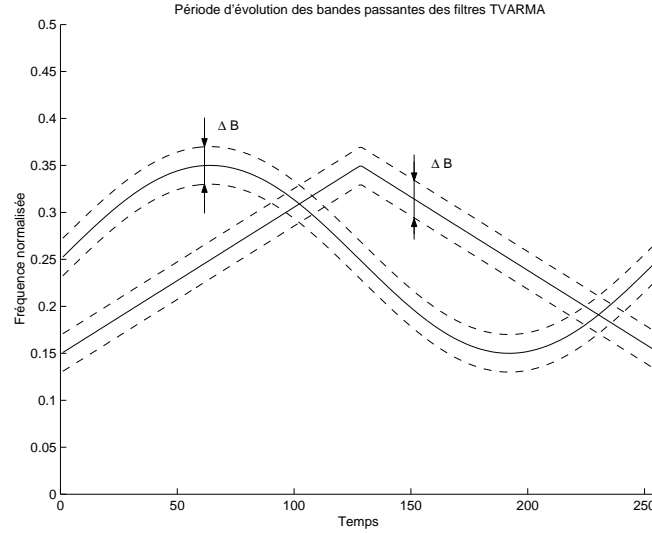


FIG. 3.14 – Période d'évolution des fréquences de coupure des 2 filtres TVARMA permettant de générer les sources. Le paramètre ΔB règle la bande passante.

Dans les deux cas nous nous plaçons en contexte déterminé ($m = n = 2$) et non-bruité, avec :

$$\mathbf{A} = \begin{bmatrix} 1 & 0.5 \\ 0.6 & 1 \end{bmatrix} \quad (3.117)$$

TFBSS est utilisé avec une distribution pseudo-Wigner-Ville lissée, $N_T = N_F = 256$ et $\varepsilon_{Grad} = 0.01$. L'ensemble des retards choisis dans SOBI est $\{1, 2, 3\}$.

Les résultats du cas non-stationnaire sont reportés sur la figure 3.15 tandis que ceux du cas stationnaire sont reportés sur la figure 3.16.

Discussion Commentons dans un premier temps le cas stationnaire. Les performances de SOBI et JADE s'améliorent lorsque la longueur des signaux augmente. Ceci est cohérent avec le fait que l'estimation des covariances et des matrices de cumulants des signaux s'améliore asymptotiquement. TFBSS donne de bons résultats, quasiment identiques à ceux de SOBI.

Dans le cas non-stationnaire, les performances de TFBSS s'améliorent lorsque T augmente. Les performances de SOBI et JADE sont relativement bonnes mais diffèrent selon la valeur de ΔB . Ainsi, dans le cas $\Delta B = 0.16$, les performances de SOBI s'améliorent lorsque T augmente. En revanche dans les deux autres cas, la longueur des signaux ne semble pas avoir une influence sur SOBI.

Ces résultats montrent encore une fois que les algorithmes SOBI et JADE, bien que construits sur des hypothèses de stationnarité des sources, peuvent donner de bons résultats sur des signaux non-stationnaires, et donc que leur champ d'action va au-delà du cas stationnaire. En revanche, il n'est pas possible de prédire d'une façon générale leur comportement asymptotique sur des signaux non-stationnaires ; chaque cas doit être étudié séparément.

Séparation de sources audio

Nous présentons maintenant les résultats de séparation d'un mélange de $n = 2$ sources audio. La source s_1 est un extrait de parole, la source s_2 est un extrait de guitare électrique saturée. Les sources sont de longueur $T = 2^{16} = 65536$, échantillonnées à la fréquence 8000 Hz (soit une durée d'environ 8s). Les sources sont représentées sur la figure 3.17.

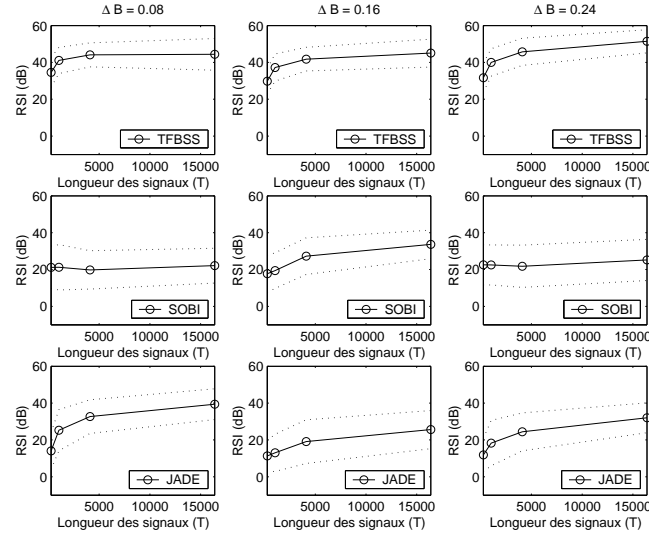


FIG. 3.15 – Comparaison de TFBSS, SOBI et JADE sur des sources non-stationnaires, en fonction de la longueur des signaux et du recouvrement spectral (en colonnes). Valeur moyenne et écart-type de \bar{RSI} sur 100 réalisations.

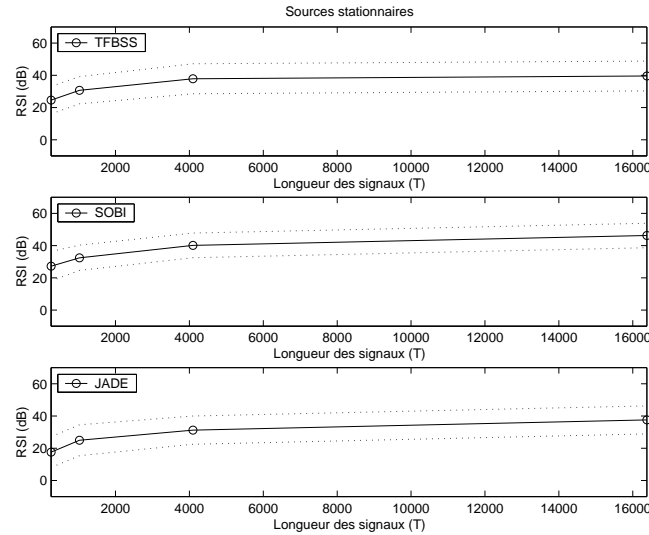


FIG. 3.16 – Comparaison de TFBSS, SOBI et JADE sur des sources stationnaires non-gaussiennes, en fonction de la longueur des signaux et du recouvrement spectral (en colonnes). Valeur moyenne et écart-type de \bar{RSI} sur 100 réalisations.

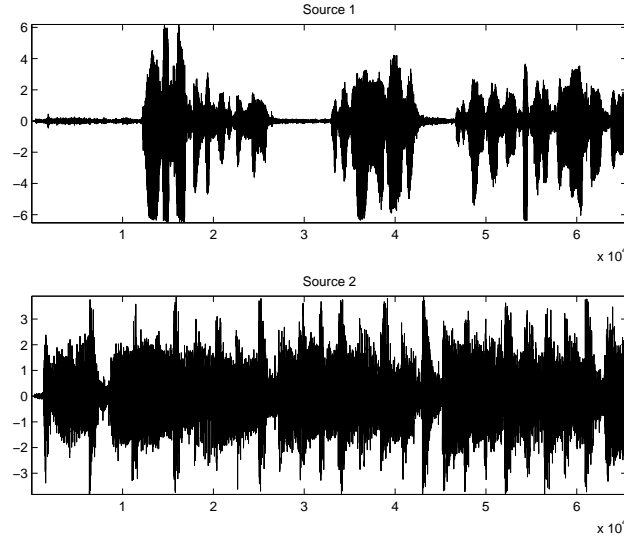


FIG. 3.17 – Représentations temporelles des sources. La source s_1 est un extrait de parole, la source s_2 est un extrait de guitare électrique saturée.

	RSI	
	\hat{s}_1	\hat{s}_2
TFBSS	31.5	63.9
BGML	43.4	33.2
SOBI	27.0	35.2
JADE	36.7	38.0

TAB. 3.1 – Résultats de séparation de TFBSS, BGML, SOBI et JADE sur un mélange de 2 sources audio.

Dans la mesure où les performances de SOBI reposent sur cette seule information, et bien que les sources audio soient non-stationnaires, nous avons représenté sur la figure 3.18 les fonctions de corrélation «empiriques» $\hat{r}_{s_1 s_2}[\tau]$, $\hat{r}_{s_1 s_1}[\tau]$ et $\hat{r}_{s_2 s_2}[\tau]$ pour $\tau \in \llbracket 0, 50 \rrbracket$.

La matrice de mélange vaut :

$$\mathbf{A} = \begin{bmatrix} 1 & 0.5 \\ 0.6 & 1 \end{bmatrix} \quad (3.118)$$

Nous avons appliqué TFBSS (avec $N_T = N_F = 256$ et $\varepsilon_{Grad} = 2/(N_T + N_F)$), BGML (avec $T_{fen} = 17$), SOBI (avec les retards $\{1, 2, 3\}$) et JADE. Les résultats sont présentés dans le tableau 3.1

Discussion Outre les résultats présentés ici, nous avons appliqué TFBSS, SOBI, JADE et BGML à de nombreux mélanges synthétiques de sources audio. Nous avons constaté que les résultats des méthodes dépendent dans une large mesure du degré de corrélation empirique des sources.

Si les sources sont bien décorréées, SOBI et JADE peuvent donner d'excellents résultats, éventuellement meilleurs que ceux fournis par TFBSS et BGML. En revanche, si la corrélation entre les sources n'est pas parfaite, TFBSS et BGML donnent généralement de meilleurs résultats. Ainsi, les résultats contenus dans le tableau 3.1 illustrent la forte sensibilité de SOBI au degré de décorrélation des signaux : bien que les sources soient *relativement* bien décorréées, TFBSS et BGML donnent de meilleurs résultats que SOBI.

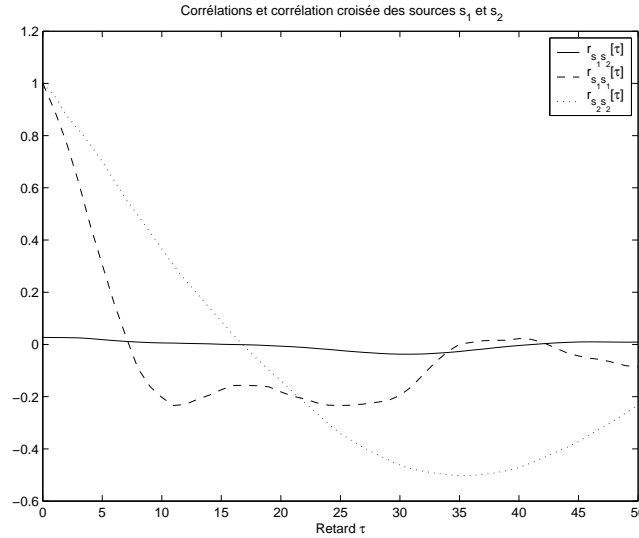


FIG. 3.18 – Fonctions de corrélation $\hat{r}_{s_1s_2}[\tau]$, $\hat{r}_{s_1s_1}[\tau]$ et $\hat{r}_{s_2s_2}[\tau]$ pour $\tau \in \llbracket 0, 50 \rrbracket$.

Bien que la longueur des signaux soit grande ($T = 65536$), on remarquera concernant TFBSS qu'il suffit de calculer les RTFs sur seulement $N_T = N_F = 256$ points temporels et fréquentiels pour obtenir de bons résultats.

Toutefois, dans la mesure où d'un point de vue auditif il est impossible de distinguer les interférences dans l'estimation d'une source au-delà d'un RSI de 30 dB, à l'écoute les quatre méthodes se valent.

3.4 Conclusion

Ce chapitre contient les principales contributions de cette thèse. Dans un premier temps nous avons proposé une relecture d'une méthode t-f existante en lui donnant une interprétation cohérente avec l'approche stochastique de la SAS présentée dans cette thèse. Puis nous avons contribué à l'amélioration des performances de cette méthode en proposant un critère de sélection des points t-f original, basé sur l'identification d'auto-termes simples des sources.

En préalable à une série de simulations, nous avons proposé des critères d'évaluation de performance originaux. Ces critères corrigent les défauts de précédents critères existant dans la littérature et, la nature de l'erreur d'estimation ayant une importance dans certaines applications, permettent de mesurer les contributions des différentes sources d'erreur dans l'estimation d'une source.

Enfin, l'ensemble des résultats en simulation présentés a montré le gain de performance significatif apporté par TFBSS sur des signaux non-stationnaires par rapport aux méthodes standard SOBI et JADE, en contexte sur-déterminé bruité. On retiendra que ces deux derniers algorithmes, bien que construits sur des hypothèses de stationnarité des sources peuvent donner d'assez bonnes performances sur des signaux non-stationnaires.

La comparaison des méthodes ATS et PHE souligne la pertinence de la recherche d'auto-termes simples des sources. La méthode ATS permet de limiter le nombre de points t-f sélectionnés en ne choisissant que les matrices $\hat{S}_{ss}^\phi[t, f]$ les plus diagonales. Au contraire, la méthode PHE sélectionne un très grand nombre de points dont une forte proportion de points de nuisance.

Bien que nécessitant un blanchiment spatial préalable, les performances de TFBSS sont, en contexte déterminé non-bruité, meilleures que celle de BGML et illustrent la pertinence de

l'approche t-f. Dans le cas déterminé non-bruité une solution optimale consisterait à élaborer une RTFS de la classe de Cohen garantissant la positivité des matrices $\hat{\mathcal{S}}_{ss}^\phi[t, f]$ afin d'utiliser comme dans BGML une procédure de diagonalisation simultanée sans contrainte orthonormale, c-à-d, sans blanchiment. L'approche adoptée par Pham dans [Pha02], exploitant une «densité spectrale instantanée» particulière des sources, peut être considérée comme un premier pas dans ce sens.

L'inconvénient majeur de TFBSS est évidemment sa complexité de calcul. Toutefois, nous avons vu, notamment sur des signaux audio, qu'il n'est pas nécessaire de calculer la RTFS des observations sur l'ensemble des points des signaux ni sur un nombre exhaustif de points fréquentiels pour obtenir de bonnes performances en des temps de calcul raisonnables.

Enfin, nous avons constaté sur des simulations qui ne sont pas présentées ici que les méthodes exposées sont très sensibles aux écarts de modélisation. Si le mélange ne s'avère pas être rigoureusement instantané, les performances de JADE, SOBI, BGML et TFBSS s'effondrent. Toutefois, ces méthodes peuvent être, dans une certaine mesure, étendues à des mélanges convolutifs à Réponse Impulsionnelle Finie. C'est cette perspective qui est étudiée dans le chapitre suivant.

Chapitre 4

Généralisation aux mélanges convolutifs RIF

Dans ce chapitre, nous montrons comment les méthodes JADE, SOBI, TFBSS et BGML exposées précédemment et basées sur une procédure de diagonalisation simultanée, peuvent être généralisées à la séparation de mélanges convolutifs à Réponse Impulsionnelle Finie (RIF) en utilisant une procédure de *bloc*-diagonalisation simultanée (BDS).

Le principe est de reformuler le mélange convolutif en un mélange linéaire instantané. Ceci est possible en considérant un «vecteur source généralisé» contenant les sources à l'instant t mais aussi *suffisamment* de versions retardées aux instants $t - 1, t - 2, \dots$ afin de prendre en compte l'effet convolutif. De la même façon, des versions retardées des observations sont ajoutées au «vecteur observation généralisé» afin de créer un mélange linéaire instantané (sur-)déterminé.

Cependant, si les sources sont mutuellement décorrélées, les composantes du vecteur source généralisé ne sont plus *toutes* décorrélées (à moins que les sources soient i.i.d). En particulier, la matrice de covariance du vecteur source généralisé est bloc-diagonale et non plus diagonale. Nous allons voir qu'il est possible de séparer les *contributions* des sources au mélange en utilisant une procédure de BDS. Toutefois, cette méthode ne fournit pas les sources originales mais *plusieurs versions filtrées de chaque source*. Une deuxième étape de *déconvolution aveugle* sera donc nécessaire pour retrouver les sources originales d'après les versions filtrées dont on dispose.

Les travaux que nous allons présenter ont été initiés par A. Belouchrani, K. Abed-Meraim et d'autres co-auteurs. L'algorithme SOBI est généralisé au cas convolutif dans [BSBAM01b], la méthode t-f [BA98] est généralisée dans [AMBL03] tandis que l'algorithme BGML est généralisé dans [BSBAM01a]. Dans ces travaux la deuxième étape de déconvolution aveugle n'est pas effectuée, bien qu'évoquée. Comme nous l'avons fait pour la séparation de mélanges linéaires instantanés, nous proposons une approche unifiée de ces travaux dans un cadre stochastique [FD03]. Ces méthodes entrent dans la catégorie des méthodes dites de *décomposition en sous-espaces* et le lecteur pourra se référer à d'autres travaux du même type dans [GL97, MJL00].

Nous allons nous intéresser plus particulièrement à l'extension au cas convolutif des méthodes SOBI et TFBSS. Cependant, les résultats que nous présentons sont encore partiels et soulignent plusieurs perspectives de recherche

D'un point de vue théorique, toute l'étude de l'identifiabilité des mélanges convolutifs RIF reste à faire. Cette dernière soulève de nombreuses difficultés dont l'issue est discutée dans le paragraphe 4.2.5.

Dans la pratique, nous avons rencontré plusieurs problèmes, encore non résolus. Le premier de ces problèmes concerne l'algorithme de BDS utilisé, décrit en annexe A.4. Si la convergence de l'algorithme est assurée par construction, sa convergence vers le minimum n'est pas garantie. Les taux d'échec sont relativement faibles sur des matrices de taille moyenne mais augmentent sensiblement avec la dimension des matrices. Ce problème est discuté plus en détail en annexe A.4.2.

Le deuxième problème concerne la sélection des points t-f dans TFBSS. La recherche d'auto-termes simples proposée dans le cas instantané n'est pas immédiatement généralisable au cas convolutif. Les résultats fournis par la procédure de sélection décrite plus loin sont encore décevants et nous soulèverons les difficultés à surmonter.

Dans la partie 4.1, nous présentons le modèle convolutif étudié et montrons comment ce dernier peut être reformulé en un mélange linéaire instantané. La partie 4.2 présente la méthode générale de SAS, puis les parties 4.3 et 4.4 présentent les implémentations de la méthode générale dans le cas stationnaire («SOBI convolutif») et non-stationnaire («TFBSS convolutif»).

4.1 Modèle et notations

Nous étudions le modèle FIR MIMO (Multi-Inputs Multi-Outputs) suivant :

$$\mathbf{x}[t] = \mathbf{H}[0] \mathbf{s}[t] + \mathbf{H}[1] \mathbf{s}[t-1] + \dots + \mathbf{H}[L] \mathbf{s}[t-L] + \mathbf{n}[t] \quad (4.1)$$

où, $\forall t \in \llbracket 0, T-1 \rrbracket$:

- $\mathbf{x}[t] = [x_1[t], \dots, x_m[t]]^T$ est le vecteur de dimension m contenant les observations,
- $\mathbf{s}[t] = [s_1[t], \dots, s_n[t]]^T$ est le vecteur de dimension n contenant les sources,
- $\forall k \in \llbracket 0, L \rrbracket$, $\mathbf{H}[k] = \{h_{ij}[k]\}$ est une matrice de dimension $m \times n$,
- $\mathbf{n}[t] = [n_1[t], \dots, n_m[t]]^T$ est un vecteur de dimension m contenant le bruit.

4.1.1 Retour à un mélange linéaire instantané

Dans ce paragraphe nous montrons comme dans [BSBAM01b, BSBAM01a, AMBL03] comment le mélange convolutif (4.1) peut être reformulé en un mélange linéaire instantané.

Soit L' un entier. On note, $\forall i \in \llbracket 1, n \rrbracket$:

$$\mathbf{S}_i[t] = [s_i[t], \dots, s_i[t - (L + L') + 1]]^T \quad (4.2)$$

et $\forall j \in \llbracket 1, m \rrbracket$

$$\mathbf{X}_j[t] = [x_j[t], \dots, x_j[t - L' + 1]]^T \quad (4.3)$$

$$\mathbf{N}_j[t] = [n_j[t], \dots, n_j[t - L' + 1]]^T \quad (4.4)$$

Par convention on pose $\mathbf{s}[-1] = \dots = \mathbf{s}[-L - L'] = \mathbf{0}$, $\mathbf{x}[-1] = \dots = \mathbf{x}[-L'] = \mathbf{0}$ et $\mathbf{n}[-1] = \dots = \mathbf{n}[-L'] = \mathbf{0}$.

On introduit :

$$\mathbf{S}[t] = [\mathbf{S}_1[t]^T, \dots, \mathbf{S}_n[t]^T]^T \quad (4.5)$$

$$\mathbf{X}[t] = [\mathbf{X}_1[t]^T, \dots, \mathbf{X}_m[t]^T]^T \quad (4.6)$$

$$\mathbf{N}[t] = [\mathbf{N}_1[t]^T, \dots, \mathbf{N}_m[t]^T]^T \quad (4.7)$$

En notant $N = n(L + L')$ et $M = mL'$, $\mathbf{S}[t]$ est un vecteur de dimension $N \times 1$, tandis que $\mathbf{X}[t]$ et $\mathbf{N}[t]$ sont des vecteurs de dimension $M \times 1$.

Pour $i \in \llbracket 1, n \rrbracket$ et $j \in \llbracket 1, m \rrbracket$, on définit la matrice de Sylvester \mathbf{A}_{ij} , de dimension $L' \times (L + L')$, par :

$$\mathbf{A}_{ij} = \begin{bmatrix} h_{ij}[0] & \dots & h_{ij}[L] & 0 & \dots & 0 \\ & \ddots & & \ddots & & \\ & & \ddots & & \ddots & \\ 0 & \dots & 0 & h_{ij}[0] & \dots & h_{ij}[L] \end{bmatrix} \quad (4.8)$$

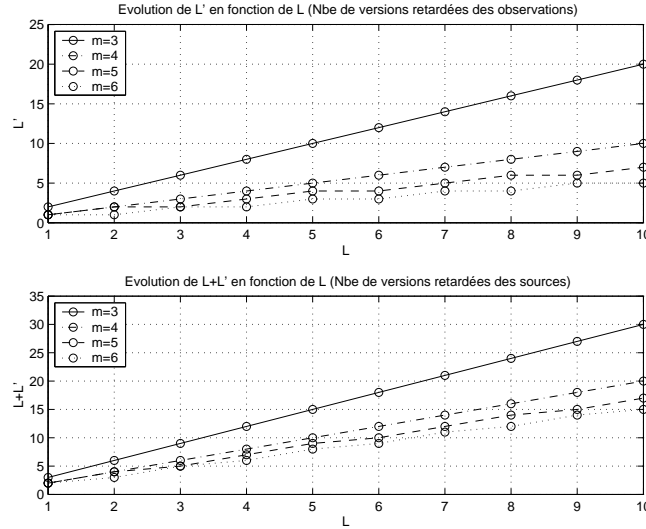


FIG. 4.1 – Evolution du nombre de versions retardées des observations (L') et des sources ($L + L'$) nécessaires pour obtenir un système (sur-)déterminé, pour $n = 2$ et plusieurs valeurs de m .

Enfin, on définit :

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \dots & \mathbf{A}_{1n} \\ \vdots & & \vdots \\ \mathbf{A}_{m1} & \dots & \mathbf{A}_{mn} \end{bmatrix} \quad (4.9)$$

\mathbf{A} est une matrice de dimensions $M \times N$ satisfaisant :

$$\mathbf{X}[t] = \mathbf{A} \mathbf{S}[t] + \mathbf{N}[t] \quad (4.10)$$

L'équation (4.10) définit un mélange linéaire instantané équivalent au mélange convolutif (4.1). Ce mélange est (sur-)déterminé si $M \geq N$, c-à-d, si L' vérifie $mL' \geq n(L + L')$, soit encore $(m - n)L' \geq nL$, ce qui nécessite $m > n$. La figure 4.1 représente l'évolution du nombre (*minimal*) de versions retardées des observations (L') et des sources ($L + L'$) nécessaires à l'obtention d'un système (sur-)déterminé, pour $n = 2$ et plusieurs valeurs de m .

4.1.2 Discussion

Nous avons montré qu'un mélange convolutif FIR peut être reformulé en un mélange linéaire instantané (sur-)déterminé (si $m > n$) en introduisant des variables adéquates.

Cependant, ce mélange est différent des mélanges linéaires instantanés étudiés aux chapitres précédents. En effet, si les sources sont modélisées par des séquences aléatoires mutuellement décorréées, les composantes de \mathbf{S} ne sont pas toutes décorréées. En effet, si les sources ne sont pas blanches, les composantes de \mathbf{S}_i sont corrélées. Cependant les séquences (vectorielles) $\mathbf{S}_1, \dots, \mathbf{S}_n$ sont mutuellement décorréées.

Ainsi, le problème de SAS décrit par l'équation (4.10) entre dans le cadre du problème de l'Analyse en Composantes Indépendantes Multidimensionnelle (ACIM), introduit par De Lathauwer dans [LCMV95] et développé par Cardoso dans [Car98b], qui propose une généralisation de l'ACI au cas où certaines des sources sont corrélées. Aussi, comme expliqué dans [Car98b] on peut dès à présent prévoir que les composantes $\mathbf{S}_1, \dots, \mathbf{S}_n$ ne pourront être estimées qu'à une matrice non-diagonale près qui modélise les indéterminations inhérentes à l'ACIM (voir le paragraphe 4.2.4).

Comme dans le cas des mélanges linéaires instantanés, on voit d'après l'équation (4.1) que les sources ne pourront *au mieux* qu'être estimées à gain et permutation près. Mais cela est-il seulement possible, et si oui à quelles conditions ? L'identifiabilité des mélanges convolutifs FIR est une tâche non triviale qui sera discutée dans la partie 4.2.5. Pour le moment nous posons les hypothèses suivantes :

- (H1) $m > n$,
- (H2) L' est choisi tel que $M \geq N$,
- (H3) Les sources s_1, \dots, s_n sont modélisées par des séquences non-stationnaires colorées et mutuellement décorréées,
- (H4) \mathbf{n} est la réalisation d'une séquence aléatoire vectorielle centrée i.i.d de covariance $\sigma^2 \mathbf{I}_m$, indépendante de \mathbf{s} .

Des hypothèses supplémentaires seront introduites au cours de l'exposé de la méthode puis récapitulées et discutées au paragraphe 4.2.5.

4.2 Méthode générale

4.2.1 Représentation temps-retard

Dans cette partie nous formulons le problème décrit par l'équation (4.10) dans le plan temps-retard.

Avec l'équation (4.10) on a :

$$\mathcal{R}_{\mathbf{X}\mathbf{X}}[t, \tau] = \mathbf{A} \mathcal{R}_{\mathbf{S}\mathbf{S}}[t, \tau] \mathbf{A}^H + \mathcal{R}_{\mathbf{N}\mathbf{N}}[t, \tau] \quad (4.11)$$

Les sources s_1, \dots, s_n étant supposées mutuellement décorréées, les séquences vectorielles $\mathbf{S}_1, \dots, \mathbf{S}_n$ sont mutuellement décorréées. La matrice de covariance $\mathcal{R}_{\mathbf{S}\mathbf{S}}[t, \tau]$ de \mathbf{S} , de dimensions $N \times N$, est donc bloc-diagonale avec n blocs de dimensions $(L + L')$:

$$\mathcal{R}_{\mathbf{S}\mathbf{S}}[t, \tau] = \begin{bmatrix} \mathcal{R}_{\mathbf{S}_1\mathbf{S}_1}[t, \tau] & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathcal{R}_{\mathbf{S}_n\mathbf{S}_n}[t, \tau] \end{bmatrix} \quad (4.12)$$

L'expression de $\mathcal{R}_{\mathbf{N}\mathbf{N}}[t, \tau]$ peut être développée davantage. Tout d'abord, le bruit étant supposé stationnaire, on a :

$$\mathcal{R}_{\mathbf{N}\mathbf{N}}[t, \tau] = \mathcal{R}_{\mathbf{N}\mathbf{N}}[\tau] \quad (4.13)$$

Les séquences de bruit n_1, \dots, n_m étant supposées mutuellement décorréées, les séquences $\mathbf{N}_1, \dots, \mathbf{N}_m$ sont mutuellement décorréées. On a donc :

$$\mathcal{R}_{\mathbf{N}\mathbf{N}}[\tau] = \begin{bmatrix} \mathcal{R}_{\mathbf{N}_1\mathbf{N}_1}[\tau] & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathcal{R}_{\mathbf{N}_m\mathbf{N}_m}[\tau] \end{bmatrix} \quad (4.14)$$

Enfin, $\forall i \in \llbracket 1, m \rrbracket$, on a :

$$\mathcal{R}_{\mathbf{N}_i\mathbf{N}_i}[\tau] = \sigma^2 \tilde{\mathbf{I}}_{L'}[\tau] \quad (4.15)$$

où $\tilde{\mathbf{I}}_{L'}[\tau]$ est la matrice de dimensions $L' \times L'$ contenant des 1 sur la τ^{eme} sous-diagonale si $0 \leq \tau < L'$ ou sur la $|\tau|^{eme}$ sur-diagonale si $-L' < \tau \leq 0$, et des zéros partout ailleurs¹. On voit ainsi que $\mathcal{R}_{\mathbf{N}\mathbf{N}}[\tau]$ ne dépend que de σ^2 .

Comme dans le cas linéaire instantané, une méthode de SAS en deux étapes (orthonormalisation et rotation) peut être dérivée de l'équation (4.10).

¹Par convention on écrira $\tilde{\mathbf{I}}_{L'}[0] = \mathbf{I}_{L'}$ et $\tilde{\mathbf{I}}_{L'}[\tau] = \mathbf{0}_{L'}$ si $|\tau| \geq L'$.

4.2.2 Orthonormalisation

La première étape consiste à trouver une matrice d'orthonormalisation. Idéalement, ce serait une matrice \mathbf{W} de dimensions $N \times M$ telle que $\mathbf{W}(\mathbf{A}\mathbf{D}\mathbf{A}^H)\mathbf{W}^H = \mathbf{I}_N$ où \mathbf{D} serait une matrice diagonale > 0 . Dans le cas convolutif, nous ne pouvons pas calculer une telle matrice. En revanche, nous pouvons calculer une matrice \mathbf{W} vérifiant :

$$\mathbf{W} \mathbf{A} \mathbf{M} \mathbf{A}^H \mathbf{W}^H = \mathbf{I}_N \quad (4.16)$$

où \mathbf{M} est une matrice bloc-diagonale strictement positive de dimensions $N \times N$ comportant n blocs de dimensions $(L + L')$. En effet, comme dans le cas linéaire instantané, une telle matrice \mathbf{W} peut être calculée d'après la matrice de corrélation des observations. Avec :

$$\mathcal{R}_{\mathbf{SS}} \stackrel{\text{def}}{=} \frac{1}{T} \sum_{t=0}^{T-1} \mathcal{R}_{\mathbf{SS}}[t, 0] \quad (4.17)$$

$$\mathcal{R}_{\mathbf{XX}} \stackrel{\text{def}}{=} \frac{1}{T} \sum_{t=0}^{T-1} \mathcal{R}_{\mathbf{XX}}[t, 0] \quad (4.18)$$

On a, compte tenu des hypothèses sur le bruit :

$$\mathcal{R}_{\mathbf{XX}} = \mathbf{A} \mathcal{R}_{\mathbf{SS}} \mathbf{A}^H + \sigma^2 \mathbf{I}_M \quad (4.19)$$

En supposant $\mathcal{R}_{\mathbf{SS}}$ et \mathbf{A} de rangs pleins, avec $\mathbf{M} = \mathcal{R}_{\mathbf{SS}}$, \mathbf{W} peut être calculée comme au paragraphe 2.2.1 d'après les vecteurs propres associées aux N valeurs propres non nulles de $\mathcal{R}_{\mathbf{XX}}$. En choisissant L' tel que M soit *strictement* supérieur à N , σ^2 peut être calculé d'après les $M - N$ plus petites valeurs propres de $\mathcal{R}_{\mathbf{XX}}$.

4.2.3 Rotation

On cherche dans un second temps la matrice orthonormale \mathbf{U} définie par :

$$\mathbf{U} = \mathbf{W} \mathbf{A} \mathcal{R}_{\mathbf{SS}}^{\frac{1}{2}} \quad (4.20)$$

On a :

$$\mathbf{W} \mathbf{A} = \mathbf{U} \mathcal{R}_{\mathbf{SS}}^{-\frac{1}{2}} \quad (4.21)$$

Soit $\underline{\mathcal{R}}_{\mathbf{XX}}[t, \tau]$ la matrice définie par :

$$\underline{\mathcal{R}}_{\mathbf{XX}}[t, \tau] = \mathbf{W} [\mathcal{R}_{\mathbf{XX}}[t, \tau] - \mathcal{R}_{\mathbf{NN}}[t, \tau]] \mathbf{W}^H \quad (4.22)$$

Avec les équations (4.11) et (4.21), on a :

$$\begin{aligned} \underline{\mathcal{R}}_{\mathbf{XX}}[t, \tau] &= \mathbf{W} \mathbf{A} \mathcal{R}_{\mathbf{SS}}[t, \tau] \mathbf{A}^H \mathbf{W}^H \\ &= \mathbf{U} (\mathcal{R}_{\mathbf{SS}}^{-\frac{1}{2}} \mathcal{R}_{\mathbf{SS}}[t, \tau] \mathcal{R}_{\mathbf{SS}}^{-\frac{H}{2}}) \mathbf{U}^H \end{aligned} \quad (4.23)$$

$\mathcal{R}_{\mathbf{SS}}^{-\frac{1}{2}}$, $\mathcal{R}_{\mathbf{SS}}^{-\frac{H}{2}}$ et $\mathcal{R}_{\mathbf{SS}}[t, \tau]$ sont des matrices bloc-diagonales de dimension $N \times N$ avec n blocs de dimensions $(L + L') \times (L + L')$. Ainsi, l'équation (4.23) montre que \mathbf{U} «bloc-diagonalise» $\underline{\mathcal{R}}_{\mathbf{XX}}[t, \tau]$, $\forall (t, \tau) \in \mathbb{Z}^2$.

En pratique, \mathbf{U} peut donc être estimée par la BDS d'un ensemble de matrices $\{\underline{\mathcal{R}}_{\mathbf{XX}}[t_i, \tau_i], i \in \llbracket 1, K \rrbracket\}$. Sous certaines hypothèses, la BDS fournit une matrice \mathbf{U}_{BDS} telle que :

$$\mathbf{U}_{BDS} = \mathbf{U} \mathbf{Q} \quad (4.24)$$

où \mathbf{Q} est une matrice unitaire de dimensions $N \times N$ qui modélise les indéterminations de la BDS. \mathbf{Q} est le produit d'une matrice bloc-diagonale avec n blocs de dimensions $(L + L') \times (L + L')$ et d'une matrice de permutation de ces blocs.

4.2.4 Estimation des sources

Dans ce paragraphe, nous calculons plusieurs estimations de chaque source à des filtres près, d'après \mathbf{U}_{BDS} et \mathbf{W} .

Soit le vecteur $\hat{\mathbf{S}}[t]$, de dimensions $N \times 1$, défini par :

$$\hat{\mathbf{S}}[t] = \mathbf{U}_{BDS}^H \mathbf{W} \mathbf{X}[t] \quad (4.25)$$

Les équations (4.24), (4.21) et (4.10) donnent :

$$\hat{\mathbf{S}}[t] = \mathbf{Q}^H \mathcal{R}_{\mathbf{SS}}^{-\frac{1}{2}} \mathbf{S}[t] + \mathbf{Q}^H \mathbf{U}^H \mathbf{W} \mathbf{N}[t] \quad (4.26)$$

En négligeant la contribution du bruit, et en notant $\mathbf{C} = \mathbf{Q}^H \mathcal{R}_{\mathbf{SS}}^{-\frac{1}{2}}$, on a :

$$\hat{\mathbf{S}}[t] \approx \mathbf{C} \mathbf{S}[t] \quad (4.27)$$

\mathbf{C} est une matrice bloc-diagonale de dimensions $N \times N$ comportant n blocs $\mathbf{C}_{11}, \dots, \mathbf{C}_{nn}$ de dimensions $(L+L') \times (L+L')$. Décomposons $\hat{\mathbf{S}}[t]$ en n sous-vecteurs $\hat{\mathbf{S}}_1[t], \dots, \hat{\mathbf{S}}_n[t]$ de dimensions $(L+L') \times 1$ tels que :

$$\hat{\mathbf{S}}[t] = [\hat{\mathbf{S}}_1[t]^T, \dots, \hat{\mathbf{S}}_n[t]^T]^T \quad (4.28)$$

Avec l'équation (4.27), $\forall i \in \llbracket 1, n \rrbracket$, on a :

$$\hat{\mathbf{S}}_i[t] \approx \mathbf{C}_{ii} \mathbf{S}_i[t] \quad (4.29)$$

Avec $\mathbf{S}_i[t] = [s_i[t], \dots, s_i[t - (L+L') + 1]]^T$, l'équation (4.29) montre donc que chaque composante de $\hat{\mathbf{S}}_i[t]$ est une version filtrée RIF de la i^{eme} source $s_i[t]$. Les coefficients du filtre sont contenus dans la ligne correspondante de \mathbf{C}_{ii} .

Ainsi, la méthode fournit $(L+L')$ versions filtrées de chaque source $s_i[t]$. Une étape de déconvolution aveugle est donc nécessaire pour estimer les sources originales d'après les versions filtrées dont on dispose. Une méthode de déconvolution aveugle basée sur une décomposition en sous-espaces est décrite en annexe B.

Remarquons que l'effet de filtrage provient de la structure bloc-diagonale (et non diagonale) de $\mathcal{R}_{\mathbf{SS}}$ et de l'indétermination de la BDS modélisée par \mathbf{Q} .

4.2.5 Discussion

Outre les hypothèses «formelles» (H1), (H2), (H3) et (H4) faites au paragraphe 4.1.2, nous avons utilisé au cours de l'exposé de la méthode générale d'autres hypothèses.

Nous avons supposé \mathbf{A} et $\mathcal{R}_{\mathbf{SS}}$ de rangs pleins lors de l'étape d'orthonormalisation, nous avons supposé que la BDS fournit une matrice de la forme $\mathbf{U} \mathbf{Q}$ où \mathbf{Q} est bloc-diagonale (à permutations de blocs près) et enfin nous avons supposé que les matrices \mathbf{C}_{ii} et les sources vérifient les conditions d'identifiabilité de la déconvolution aveugle (voir annexe B.4).

Comment traduire l'ensemble de ces hypothèses en termes de caractéristiques des sources et des coefficients de filtres des matrices $\mathbf{H}[k]$? C'est tout l'enjeu de l'étude de l'identifiabilité des mélanges convolutifs FIR. Les hypothèses faites sur \mathbf{A} et $\mathcal{R}_{\mathbf{SS}}$ peuvent être rapprochées des conditions d'identifiabilité de l'identification aveugle d'un système SIMO [AMQH97]. L'étude des conditions d'unicité à blocs près de la BDS est un problème entier mais abordable. En revanche, traduire les conditions d'identifiabilité de la déconvolution aveugle de $\hat{\mathbf{S}}_i$, $\forall i \in \llbracket 1, n \rrbracket$ relève du réel casse-tête...

En pratique nous n'avons que rarement rencontré de problème d'identifiabilité. Il nous est arrivé d'obtenir une matrice \mathbf{U}_{bds} dans une forme autre que celle escomptée mais généralement l'ajout d'une matrice à l'ensemble des matrices à bloc-diagonaliser a réglé le problème.

Nous allons voir que la méthode SOBI généralisée au cas convolutif donne d'excellents résultats sur des mélanges synthétiques aléatoires de sources audio. Toutefois, nous sommes conscients que toute l'étude de l'identifiabilité du problème reste à faire.

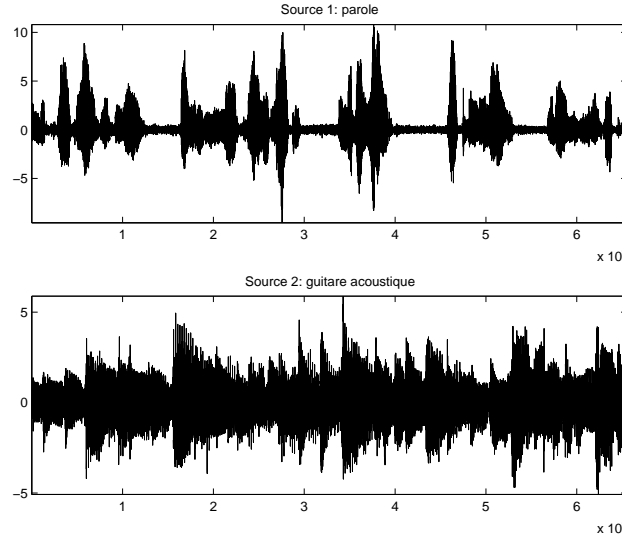


FIG. 4.2 – Représentations temporelles des sources. La source s_1 est un extrait de parole, la source s_2 est un extrait de guitare acoustique.

4.3 SOBI convolutif

Si les sources sont modélisées par des séquences stationnaires, on a tout simplement :

$$\mathcal{R}_{\mathbf{S}\mathbf{S}}[t, \tau] = \mathcal{R}_{\mathbf{S}\mathbf{S}}[\tau] \quad (4.30)$$

$$\mathcal{R}_{\mathbf{X}\mathbf{X}}[t, \tau] = \mathcal{R}_{\mathbf{X}\mathbf{X}}[\tau] \quad (4.31)$$

Si les signaux sont supposés ergodiques, une estimation de $\mathcal{R}_{\mathbf{X}\mathbf{X}}[\tau]$ est :

$$\hat{\mathcal{R}}_{\mathbf{X}\mathbf{X}}[\tau] = \frac{1}{T - \tau} \sum_{t=0}^{T-\tau-1} \mathbf{X}[t + \tau] \mathbf{X}[t]^H \quad (4.32)$$

Nous présentons quelques résultats fournis par SOBI convolutif sur plusieurs mélanges synthétiques de $n = 2$ sources audio. La source s_1 est un extrait de parole, la source s_2 est un extrait de guitare acoustique. Les sources sont de longueur $T = 2^{16} = 65536$, échantillonnées à la fréquence 8000 Hz (soit une durée d'environ 8s).

Les sources sont représentées sur la figure 4.2. La figure 4.3 représente les fonctions de corrélation empiriques $\hat{r}_{s_1 s_2}[\tau]$, $\hat{r}_{s_1 s_1}[\tau]$ et $\hat{r}_{s_2 s_2}[\tau]$ pour $\tau \in \llbracket 0, 50 \rrbracket$. La courbe de $\hat{r}_{s_1 s_2}[\tau]$ montre que les sources ne sont quasiment pas corrélées.

Nous avons appliqué la méthode à des mélanges non bruités choisis aléatoirement, pour $L \in \{1, 3, 5\}$ et $m \in \{3, 4, 5, 6\}$. Les retards $\{1, 2, 3, 4\}$ ont été utilisés pour le calcul des matrices de covariance à bloc-diagonaliser. Après l'étape de BDS nous avons calculé le critère de bloc-diagonalité suivant de la matrice \mathbf{C} (après correction des éventuelles permutations de blocs) :

$$C_{bd}(\mathbf{C}) = \frac{\sum_{i=1}^n \sum |\mathbf{C}_{ii}|}{\sum |\mathbf{C}|} \quad (4.33)$$

où $\sum |\mathbf{M}|$ désigne la somme des valeurs absolues des coefficients d'une matrice quelconque \mathbf{M} . Le calcul de $C_{bd}(\mathbf{C})$ permet de mesurer la qualité de la séparation des contributions des sources dans le mélange. Puis, après l'étape de déconvolution aveugle de chaque source, nous avons calculé les RSD, RSI et RSA des estimations \hat{s}_1 et \hat{s}_2 . Nous avons aussi mesuré les durées de calcul requises dans chaque cas (sur un ordinateur Pentium 3 cadencé à 733 MHz). La méthode

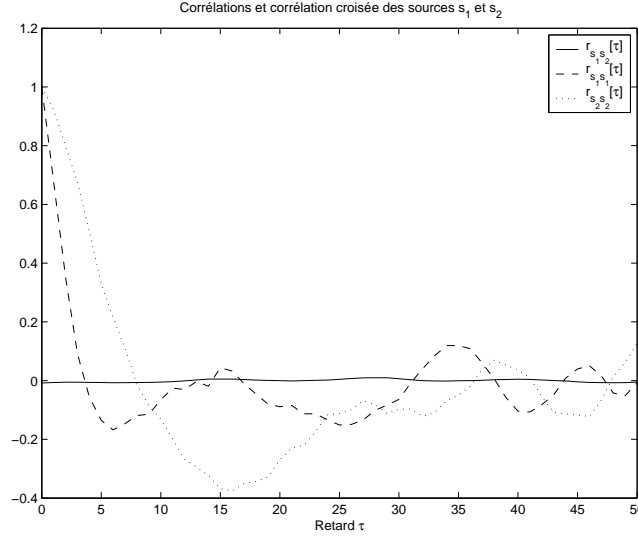


FIG. 4.3 – Fonctions de corrélation $\hat{r}_{s_1 s_2}[\tau]$, $\hat{r}_{s_1 s_1}[\tau]$ et $\hat{r}_{s_2 s_2}[\tau]$ pour $\tau \in \llbracket 0, 50 \rrbracket$.

	L=1						L=2						L=3					
	RSD		RSI		RSA		RSD		RSI		RSA		RSD		RSI		RSA	
	\hat{s}_1	\hat{s}_2	\hat{s}_1	\hat{s}_2	\hat{s}_1	\hat{s}_2	\hat{s}_1	\hat{s}_2	\hat{s}_1	\hat{s}_2	\hat{s}_1	\hat{s}_2	\hat{s}_1	\hat{s}_2	\hat{s}_1	\hat{s}_2	\hat{s}_1	\hat{s}_2
m=3	41	55	42	92	49	55	40	39	43	59	43	39	39	33	44	57	41	33
m=4	41	58	42	83	50	58	42	49	43	58	46	49	39	36	42	69	43	36
m=5	41	58	42	83	50	58	41	51	43	63	47	51	40	39	43	59	43	39
m=6	41	58	42	83	50	58	41	51	43	64	47	51	41	42	44	57	43	43

TAB. 4.1 – Résultats de SOBI convolutif pour différents paramètres de mélange, pour $n = 2$ sources. Valeurs en dB des RSD, RSI et RSA des estimations des sources.

de BDS utilisée est la méthode M3 décrite en annexe A.4.2. La méthode de déconvolution aveugle utilisée est celle décrite en annexe B, avec $F = L + L'$. Les résultats sont présentés dans les tableaux 4.1, 4.2 et 4.3.

Discussion Les valeurs de $C_{bd}(\mathbf{C})$ présentées dans le tableau 4.2 montrent que la première étape de la méthode, c-à-d, la séparation des contributions des sources dans le mélange, donne d'excellents résultats : les valeurs de $C_{bd}(\mathbf{C})$ sont toutes supérieures à 0.99.

L'évolution des valeurs des RSD, RSI et RSA des estimations des sources dans le tableau 4.1 est donc essentiellement due à la deuxième étape de déconvolution aveugle. A L fixé les résultats s'améliorent lorsque m augmente car $L + L'$ diminue, ce qui signifie que le nombre de coefficients de filtres à estimer diminue. Pour L fixé, les résultats ne changent plus au-delà d'une certaine valeur de m car $L + L'$ reste constant (voir la figure 4.1). On remarquera que la méthode générale

	L=1	L=3	L=5
m=3	0.9952	0.9940	0.9911
m=4	0.9955	0.9961	0.9929
m=5	0.9955	0.9952	0.9940
m=6	0.9955	0.9952	0.9948

TAB. 4.2 – Résultats de SOBI convolutif pour différents paramètres de mélange, pour $n = 2$ sources. Valeurs de $C_{bd}(\mathbf{C})$.

	L=1	L=3	L=5
m=3	0.0467	2.0917	45.6054
m=4	0.0199	0.3355	2.4272
m=5	0.0245	0.1659	2.1912
m=6	0.0281	0.1754	1.0277

TAB. 4.3 – Résultats de SOBI convolutif pour différents paramètres de mélange, pour $n = 2$ sources. Durées de calcul en minutes.

introduit des artefacts sur les sources, dus aux effets de filtrage.

D'un point de vue auditif, puisque les RSI et RSA des estimations sont dans chaque cas supérieurs à 30 dB, les estimations semblent parfaites à l'écoute.

Le tableau 4.3 montre que l'augmentation du nombre d'observations permet évidemment de diminuer les temps de calcul, qui correspondent dans une très large mesure à la durée de BDS. Encore une fois, ceci est du au fait que la taille des blocs $L + L'$ des matrices $\mathcal{R}_{\mathbf{SS}}[\tau]$ décroît lorsque m augmente.

Enfin, nous avons constaté que le choix des coefficients de mélange n'a pas d'influence significative sur les résultats de séparation, ce qui tendrait à penser que la méthode possède la propriété d'équivariance. Ceci reste cependant à démontrer.

4.4 Quelle implémentation pour des sources non-stationnaires ?

Si les sources sont localement stationnaires, une estimation locale des covariances peut être réalisée, ce qui revient à généraliser BGML comme dans [BSBAM01a]. Encouragés par les bonnes performances de l'approche t-f sur des signaux non-stationnaires dans le cas linéaire instantané, nous avons souhaité généraliser TFBSS au cas convolutif. Si la généralisation de la méthode en théorie ne présente pas de difficulté, dans la pratique la sélection des point t-f est un problème. Nous proposons quelques éléments de recherche pour la généralisation de l'approche t-f au cas convolutif mais nous verrons qu'un critère de sélection robuste des points t-f reste à déterminer.

4.4.1 L'approche temps-fréquence

Dans le plan t-f, l'équation (4.11) devient :

$$\mathcal{S}_{\mathbf{XX}}[t, f] = \mathbf{A} \mathcal{S}_{\mathbf{SS}}[t, f] \mathbf{A}^H + \mathcal{S}_{\mathbf{NN}}[t, f] \quad (4.34)$$

On définit $\underline{\mathcal{S}}_{\mathbf{XX}}[t, f]$ par :

$$\underline{\mathcal{S}}_{\mathbf{XX}}[t, f] = \mathbf{W} [\mathcal{S}_{\mathbf{XX}}[t, f] - \mathcal{S}_{\mathbf{NN}}[t, f]] \mathbf{W}^H \quad (4.35)$$

On a alors :

$$\underline{\mathcal{S}}_{\mathbf{XX}}[t, f] = \mathbf{U} (\mathcal{R}_{\mathbf{SS}}^{-\frac{1}{2}} \mathcal{S}_{\mathbf{SS}}[t, f] \mathcal{R}_{\mathbf{SS}}^{-\frac{H}{2}}) \mathbf{U}^H \quad (4.36)$$

Les matrices $\mathcal{R}_{\mathbf{SS}}^{-\frac{1}{2}}$, $\mathcal{R}_{\mathbf{SS}}^{-\frac{H}{2}}$ et $\mathcal{S}_{\mathbf{SS}}[t, f]$ étant bloc-diagonales, \mathbf{U} peut donc être estimée par BDS d'un ensemble de matrices $\{\underline{\mathcal{S}}_{\mathbf{XX}}[t_i, f_i], i = 1 \dots K\}$ plutôt qu'un ensemble de matrices $\{\mathcal{R}_{\mathbf{XX}}[t_i, \tau_i], i = 1 \dots K\}$ ².

²Dans le cas où la matrice \mathbf{U} est réelle, on remarque que \mathbf{U} diagonalise simultanément $\text{Re}\{\underline{\mathcal{S}}_{\mathbf{XX}}[t, f]\}$ et $\text{Im}\{\underline{\mathcal{S}}_{\mathbf{XX}}[t, f]\}$. Ne disposant que d'un algorithme de BDS de matrices réelles nous avons utilisé cette dernière astuce en pratique pour estimer \mathbf{U} , dans le cas réel.

En pratique, des estimations $\hat{\mathbf{W}}$ et $\hat{\sigma}^2$ sont calculées d'après les éléments propres de :

$$\hat{\mathcal{R}}_{\mathbf{X}\mathbf{X}} \stackrel{\text{def}}{=} \frac{1}{T} \sum_{t=0}^{T-1} \mathbf{X}[t] \mathbf{X}[t]^H \quad (4.37)$$

$$\approx \mathbf{A} \hat{\mathcal{R}}_{\mathbf{S}\mathbf{S}} \mathbf{A}^H + \sigma^2 \mathbf{I}_M \quad (4.38)$$

Comme dans le paragraphe 3.2.2, $\mathcal{S}_{\mathbf{X}\mathbf{X}}[t, f]$ est estimé par $\hat{\mathcal{S}}_{\mathbf{X}\mathbf{X}}^\phi[t, f]$. On a alors :

$$\hat{\mathcal{S}}_{\mathbf{X}\mathbf{X}}^\phi[t, f] \approx \mathbf{U}(\hat{\mathcal{R}}_{\mathbf{S}\mathbf{S}}^{-\frac{1}{2}} \hat{\mathcal{S}}_{\mathbf{S}\mathbf{S}}^\phi[t, f] \hat{\mathcal{R}}_{\mathbf{S}\mathbf{S}}^{-\frac{H}{2}}) \mathbf{U}^H \quad (4.39)$$

avec

$$\hat{\mathcal{S}}_{\mathbf{X}\mathbf{X}}^\phi[t, f] \stackrel{\text{def}}{=} \hat{\mathbf{W}} \left(\hat{\mathcal{S}}_{\mathbf{X}\mathbf{X}}^\phi[t, f] - \hat{\mathcal{S}}_{\mathbf{N}\mathbf{N}}^\phi[t, f] \right) \hat{\mathbf{W}}^H \quad (4.40)$$

où $\hat{\mathcal{S}}_{\mathbf{N}\mathbf{N}}^\phi[t, f]$ désigne l'estimation de $\mathcal{S}_{\mathbf{N}\mathbf{N}}[t, f]$, calculée d'après l'expression (4.14) de $\mathcal{R}_{\mathbf{N}\mathbf{N}}[\tau]$ mais où σ est remplacé par $\hat{\sigma}$.

4.4.2 Sélection des points t-f

Comme dans le cas instantané, il est en pratique nécessaire d'identifier des points t-f pour lesquels la matrices $\hat{\mathcal{S}}_{\mathbf{S}\mathbf{S}}^\phi[t, f]$ est bloc-diagonale. Nous avons essayé de généraliser la recherche d'auto-termes simples dans le cas instantané à la recherche de matrices $\hat{\mathcal{S}}_{\mathbf{S}\mathbf{S}}^\phi[t, f]$ bloc-diagonales avec un seul bloc non-nul.

Dans le cas instantané nous sélectionnions les matrices $\hat{\mathcal{S}}_{\mathbf{X}\mathbf{X}}^\phi[t, f]$ de rang proche de 1, dans le cas convolutif nous avons donc essayé de sélectionner les matrices $\hat{\mathcal{S}}_{\mathbf{X}\mathbf{X}}^\phi[t, f]$ de rang $L + L'$, c-à-d, de sélectionner les maxima locaux du critère :

$$C_3[t, f] = \frac{\sum_{i=1}^{L+L'} \lambda_i[t, f]}{\sum |\text{eig}(\hat{\mathcal{S}}_{\mathbf{X}\mathbf{X}}^\phi[t, f])|} \quad (4.41)$$

où $\lambda_1[t, f], \dots, \lambda_{(L+L')}[t, f]$ désignent les $L + L'$ plus grandes valeurs propres en valeur absolue de $\hat{\mathcal{S}}_{\mathbf{X}\mathbf{X}}^\phi[t, f]$. Nous ferons référence à cette méthode par l'abréviation BATS (pour «Bloc Auto-terme Simple»).

Nous allons tester la validité de ce critère sur plusieurs signaux. Dans chaque cas le critère de bloc-diagonalité de $\hat{\mathcal{S}}_{\mathbf{S}\mathbf{S}}^\phi[t, f]$ suivant est calculé sur tout le plan t-f :

$$C_{bd}(\hat{\mathcal{S}}_{\mathbf{S}\mathbf{S}}^\phi[t, f]) = \frac{\sum_{i=1}^n \sum |\hat{\mathcal{S}}_{\mathbf{S}_i \mathbf{S}_i}^\phi[t, f]|}{\sum |\hat{\mathcal{S}}_{\mathbf{S}\mathbf{S}}^\phi[t, f]|} \quad (4.42)$$

Afin de vérifier si les matrices $\hat{\mathcal{S}}_{\mathbf{S}\mathbf{S}}^\phi[t, f]$ bloc-diagonales n'ont bien qu'un seul bloc non-nul nous calculons aussi le critère suivant sur tout le plan t-f :

$$C_{1bd}(\hat{\mathcal{S}}_{\mathbf{S}\mathbf{S}}^\phi[t, f]) = \frac{\max_i \sum |\hat{\mathcal{S}}_{\mathbf{S}_i \mathbf{S}_i}^\phi[t, f]|}{\sum |\hat{\mathcal{S}}_{\mathbf{S}\mathbf{S}}^\phi[t, f]|} \quad (4.43)$$

$C_{1bd}(\hat{\mathcal{S}}_{\mathbf{S}\mathbf{S}}^\phi[t, f])$ est proche de 1 si $\hat{\mathcal{S}}_{\mathbf{S}\mathbf{S}}^\phi[t, f]$ est bloc-diagonale avec un bloc dominant.

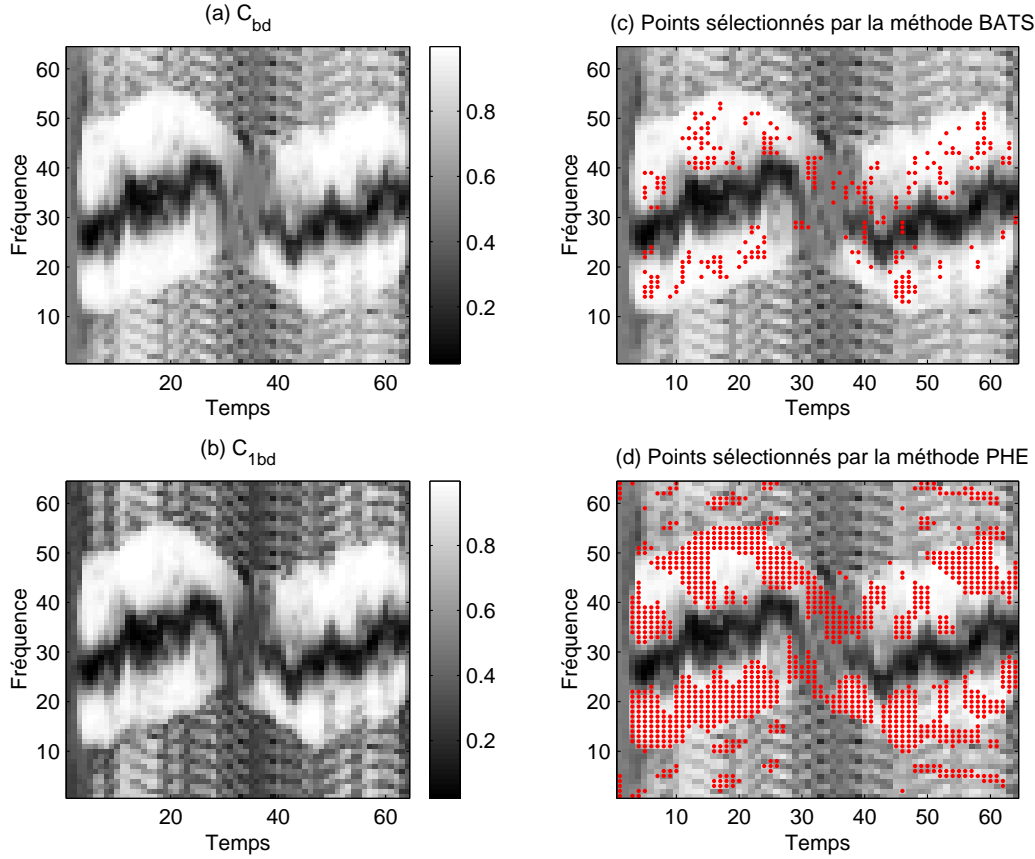


FIG. 4.4 – Localisation des points t-f sélectionnés, pour un mélange de $n = 2$ sources TVARMA de longueur $T = 256$, avec $m = 3$ et $L = 2$. (a) Critère de bloc-diagonalité $C_{bd}(\hat{\mathcal{S}}_{\mathbf{SS}}^\phi[t, f])$, (b) Critère de bloc-diagonalité avec un bloc dominant $C_{1bd}(\hat{\mathcal{S}}_{\mathbf{SS}}^\phi[t, f])$, (c) Points sélectionnés par la méthode BATS (en rouge), superposés à $C_{bd}(\hat{\mathcal{S}}_{\mathbf{SS}}^\phi[t, f])$, (d) Points sélectionnés par la méthode PHE (en rouge), superposés à $C_{1bd}(\hat{\mathcal{S}}_{\mathbf{SS}}^\phi[t, f])$.

Sources TVARMA Nous avons créé un mélange convolutif ($m = 3$, $L = 2$) de $n = 2$ sources TVARMA (comme au paragraphe 3.3), de longueur $T = 256$, l'une de profil t-f linéaire, l'autre de profil sinusoïdal, avec $\Delta B = 0.04$.

Les figures 4.4 (a) et (b) représentent $C_{bd}(\hat{\mathcal{S}}_{\mathbf{SS}}^\phi[t, f])$ et $C_{1bd}(\hat{\mathcal{S}}_{\mathbf{SS}}^\phi[t, f])$, pour $N_T = N_F = 64$. Les deux critères C_{bd} et C_{1bd} sont quasiment identiques, ce qui montre que les matrices $\hat{\mathcal{S}}_{\mathbf{SS}}^\phi[t, f]$ bloc-diagonales n'ont effectivement qu'un seul bloc non-nul.

La figure 4.4 (c) montre le résultat de la sélection des points t-f correspondant aux maxima locaux de C_3 , en utilisant une procédure de gradient avec $\varepsilon_{Grad} = 2/(N_T + N_F)$. Les points rouges représentent les points sélectionnés, superposés au tracé de $C_{bd}(\hat{\mathcal{S}}_{\mathbf{SS}}^\phi[t, f])$. La figure 4.4 (d) montre le résultat de sélection de la méthode PHE décrite dans le paragraphe 3.2.3.

Comme dans le cas linéaire instantané, la méthode PHE sélectionne de nombreux points, dont une très large proportion de points de nuisance. La méthode BATS sélectionne moins de points, mais contrairement au cas linéaire instantané, elle sélectionne aussi une proportion non négligeable de points de nuisance, proportion suffisamment importante pour affecter les résultats de séparation.

En effet, la méthode t-f appliquée au mélange avec la méthode BATS fournit seulement $C_{bd}(\mathbf{C}) = 0.956$, valeur légèrement supérieure à celle fournit par SOBI convolutif égale à 0.931. Après déconvolution aveugle, les RSD des estimations des sources fournies par les deux méthodes sont semblables (tableau 4.4).

	RSD	
	\hat{s}_1	\hat{s}_2
TFBSS convolutif	18.6	18.6
SOBI convolutif	19.2	16.4

TAB. 4.4 – Résultats de SOBI convolutif et TFBSS convolutif sur un mélange de 2 sources TVARMA.

Sources audio Appliquée aux signaux audio utilisés dans le paragraphe 4.3, les résultats de la méthode t-f empirient. La figure 4.5 montre les résultats de sélection des méthodes BATS et PHE sur un mélange comportant $m = 3$ observations avec des filtres de longueur $L = 2$, avec $N_T = N_F = 64$. La proportion de points de nuisance sélectionnés par les deux méthodes augmentent et les résultats de TFBSS convolutif sont très mauvais. En effet, on obtient seulement $C_{bd}(\mathbf{C}) = 0.701$, alors que SOBI fournit l'excellente valeur de 0.9961.

Discussion Le cas convolutif révèle les limites du critère de sélection basé sur l'identification de (bloc) auto-termes simples de sources. Dans le cas linéaire instantané (paragraphe 3.2.3), nous avons souligné que la méthode ATS sélectionne *toutes* les matrices $\hat{\mathcal{S}}_{ss}^\phi[t, f]$ de rang proche de 1. En pratique, nous avons constaté que toutes les matrices $\hat{\mathcal{S}}_{ss}^\phi[t, f]$ de rang 1 correspondaient à des auto-termes simples des sources, donc à des matrices diagonales.

En revanche, ce principe ne fonctionne plus dans le cas convolutif. Il apparaît que les matrices $\hat{\mathcal{S}}_{ss}^\phi[t, f]$ de rang $L + L'$ ne correspondent pas toutes à des matrices bloc-diagonales avec un seul bloc, et ceci explique les mauvaises performances de la méthode.

Cependant, les figures 4.4 et 4.5 montrent que l'hypothèse que les matrices $\hat{\mathcal{S}}_{ss}^\phi[t, f]$ bloc-diagonales n'ont qu'un seul bloc tient toujours, et que c'est donc résolument ce type de matrices $\hat{\mathcal{S}}_{ss}^\phi[t, f]$ qu'il faut rechercher. L'idée que les points t-f correspondant à des blocs auto-termes simples ne sont pas isolés mais appartiennent à des zones t-f (à la manière de [ADW01]) pourrait servir de piste à l'élaboration d'un meilleur critère.

4.5 Conclusion

Bien que plusieurs difficultés restent à surmonter - l'identifiabilité théorique des mélanges convolutifs FIR, la convergence de l'algorithme de BDS, la sélection des points dans le cas t-f -, les résultats obtenus par SOBI convolutif sur des mélanges courts de sources bien décorréliées sont très encourageants.

Le net avantage de l'approche que nous avons décrite est qu'elle permet une séparation *complète* des mélanges convolutifs RIF, contrairement à la majorité des méthodes existantes qui se limitent à une séparation partielle, c-à-d, à l'estimation d'une *seule* version filtrée de chaque source [WFO93, SS02, CRG95] (l'étape de déconvolution aveugle est alors impossible).

En contrepartie, la complexité de calcul nécessaire à la séparation complète est beaucoup plus importante et ne permet pas en l'état des travaux de s'attaquer à la plupart des mélanges réels, par exemple audio. En effet, considérons par exemple la séparation de $n = 2$ sources audio d'après $m = 4$ microphones (par exemple 2 micros directs, 2 micros d'ambiance) dans une salle possédant un temps de réverbération moyen de 1s. Si les signaux sont échantillonnés à (seulement) 8000 Hz, cela signifie qu'il faut envisager la BDS de matrices comportant 2 blocs diagonaux de dimensions 16000×16000 ...

La procédure de BDS est la clé de la méthode, et c'est avant tout sur ce point que les efforts doivent porter pour envisager d'améliorer les capacités de la méthode.

Remarquons enfin qu'un autre inconvénient de la méthode réside dans le fait qu'il est nécessaire de connaître la longueur L des filtres du mélange. Il n'est pas possible de surestimer cette longueur car la matrice \mathbf{A} n'est alors plus de rang plein. En revanche on peut envisager de

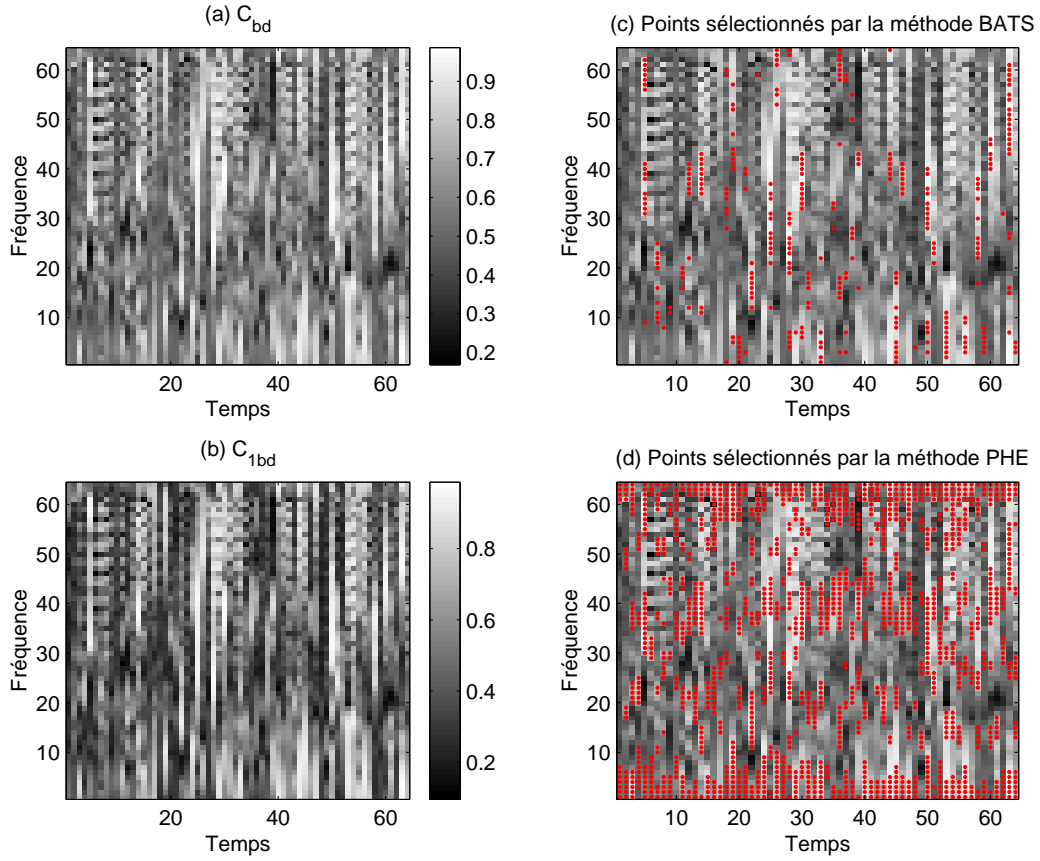


FIG. 4.5 – Localisation des points t-f sélectionnés pour un mélange de $n = 2$ sources audio de longueur $T = 2^{16}$, avec $m = 3$ et $L = 2$. (a) Critère de bloc-diagonalité $C_{bd}(\hat{\mathcal{S}}_{\mathbf{SS}}^\phi[t, f])$, (b) Critère de bloc-diagonalité avec un bloc dominant $C_{1bd}(\hat{\mathcal{S}}_{\mathbf{SS}}^\phi[t, f])$, (c) Points sélectionnés par la méthode BATS (en rouge), superposés à $C_{bd}(\hat{\mathcal{S}}_{\mathbf{SS}}^\phi[t, f])$, (d) Points sélectionnés par la méthode PHE (en rouge), superposés à $C_{bd}(\hat{\mathcal{S}}_{\mathbf{SS}}^\phi[t, f])$.

sous-estimer cette longueur, ce qui dans le cas d'enregistrements audio en salle revient à estimer les L premiers coefficients significatifs des réponses impulsionnelles de la salle entre les émetteurs et les micros d'enregistrement.

Conclusion générale

Dans les chapitres 1, 2 et 3 de cette thèse nous avons présenté une approche unifiée dans un cadre stochastique de plusieurs méthodes de séparation aveugle de mélanges linéaires instantanés.

Dans le chapitre 1 nous avons présenté un cadre pour l'étude de l'identifiabilité de ces mélanges. Puis, nous nous sommes intéressés à quatre modèles standard de sources identifiables. Ces modèles reposent sur l'indépendance mutuelle des séquences aléatoires modélisant les sources, indépendance à l'ordre 2 lorsque les sources possèdent une diversité *temporelle et/ou fréquentielle*, indépendance stricte lorsque les sources sont modélisées par des séquences i.i.d. Puis, dans les chapitres 2 et 3 nous avons présenté plusieurs méthodes de SAS construites sur les hypothèses de ces modèles. Nous avons souhaité souligner la distinction entre les hypothèses garantissant l'identifiabilité du problème (la notion d'indépendance mutuelle) des hypothèses «de travail» utilisées pour construire des estimateurs de la matrice de mélange.

Si nous ne nous sommes intéressés qu'à des modèles basés sur la propriété d'indépendance mutuelle des sources, d'autres types de modèles existent néanmoins, et le cadre d'étude de l'identifiabilité que nous avons proposé se prête bien sûr tout aussi bien à leur analyse. Ainsi, alternativement à l'utilisation de la notion d'indépendance, certaines méthodes de SAS reposent sur l'utilisation de la propriété de *parcimonie* des sources. Cette propriété signifie que les sources peuvent être représentées par un nombre restreint de coefficients dans une base de décomposition adaptée. Par le fait du mélange les observations deviennent moins parcimonieuses et l'objectif des méthodes de séparation est alors de restaurer la propriété de parcimonie [ZPBK01, Gri02]. D'autres méthodes encore reposent sur la positivité des sources [BDBG03, OP03]. L'identifiabilité des modèles de sources associés à ces méthodes mérite d'être étudiée en détails dans le cadre que nous avons proposé.

L'approche de la SAS que nous avons adoptée est une approche stochastique. Notre démarche a consisté à présenter des méthodes théoriques utilisant certaines statistiques des sources puis à utiliser en pratique des estimateurs de ces sources. Selon la nature des sources certaines statistiques sont plus adaptées que d'autres et donnent de meilleurs résultats. Ainsi, nous avons vu que, parmi les méthodes présentées, l'algorithme TFBSS est l'algorithme le mieux adapté aux sources non-stationnaires car il prend en compte la non-stationnarité des sources. Toutefois, bien qu'ils soient construits sur l'hypothèse de stationnarité des sources, nous avons pu constater que JADE et SOBI peuvent donner de bons résultats sur des sources non-stationnaires, ce qui laisse présager que leur champ d'action s'étend au-delà des sources stationnaires. Ceci signifie que les estimateurs de cumulants d'ordre 4 et de covariances, appliqués à des signaux non-stationnaires, conservent un sens qu'il convient de déterminer.

Si l'utilisation d'un contexte stochastique est une habitude courante en matière de SAS³, nous pensons qu'il serait intéressant de reformuler les méthodes présentées précédemment dans un contexte déterministe en rapprochant la notion de décorrélation de la notion d'orthogonalité dans l'espace de Hilbert des signaux d'énergie finie⁴ (c-à-d, dans le contexte utilisé pour

³Les méthodes basées sur la parcimonie ou la positivité des sources commencent à changer cette habitude.

⁴Merci à Jérôme Idier de nous en avoir inspiré l'idée.

l'élaboration des critères d'évaluation présentés dans le paragraphe 3.3.1). Les performances des estimateurs de la matrice de mélange obtenus seraient étudiées en formulant *alors* des hypothèses stochastiques sur les sources afin de calculer le biais et la variance de ces estimateurs.

Enfin, dans le chapitre 4 nous avons évoqué une manière d'appréhender le problème de la séparation des mélanges convolutifs RIF. Si de nombreuses difficultés restent à élucider, nous avons souhaité montrer que la séparation *complète* des mélanges convolutifs est chose possible. D'un point de vue pratique le problème majeur à résoudre concerne la convergence de l'algorithme de BDS proposé. Dans un deuxième temps, il serait souhaitable de trouver un critère efficace de sélection des points t-f afin de retrouver sur des mélanges convolutifs les bonnes propriétés de TFBSS pour des mélanges instantanés. D'un point de vue théorique, toute l'étude de l'identifiabilité du problème reste à faire.

Comme nous pouvons le constater, beaucoup de chemin reste à parcourir avant de pouvoir envisager le problème complet de séparation de sources évoqué dans l'introduction de cette thèse, à savoir le problème standard de la séparation de musiciens ou locuteurs se déplaçant dans une salle. Toutefois, l'intérêt croissant porté à la SAS pousse à l'optimisme. Et n'oublions pas que la séparation de sources est encore jeune, tout juste majeure cette année...

Annexe A

Diagonalisation et bloc-diagonalisation simultanée de matrices

A.1 Diagonalisation approchée

A.1.1 Base orthonormale complexe

Nous présentons dans ce paragraphe une méthode de type Jacobi pour la *diagonalisation approchée* d'une matrice $\mathbf{A} \in \mathbb{C}^{n \times n}$ dans une base orthonormale complexe, méthode qui sera généralisée à la diagonalisation simultanée de matrices dans le prochain paragraphe.

Si \mathbf{A} est diagonalisable dans une base orthonormale, alors l'algorithme que nous allons présenter calcule une telle base. Si \mathbf{A} n'est pas diagonalisable dans une base orthonormale, alors l'algorithme produit une base orthonormale dans laquelle \mathbf{A} est *la plus diagonale possible*, c-à-d que l'on recherche une matrice orthonormale $\mathbf{U} \in \mathbb{C}^{n \times n}$ telle que

$$\mathbf{U}\mathbf{A}\mathbf{U}^H = \mathbf{D} \quad (\text{A.1})$$

où $\mathbf{D} \in \mathbb{C}^{n \times n}$ minimise le critère de diagonalité suivant :

$$\text{off}(\mathbf{M}) \stackrel{\text{def}}{=} \sum_{1 \leq i \neq j \leq n} |m_{ij}|^2 \quad (\text{A.2})$$

On recherche donc \mathbf{U} par minimisation du critère suivant :

$$C_d(\mathbf{V}; \mathbf{A}) \stackrel{\text{def}}{=} \text{off}(\mathbf{V} \mathbf{A} \mathbf{V}^H) \quad (\text{A.3})$$

où $\mathbf{V} \in \mathbb{C}^{n \times n}$, sous la contrainte d'orthonormalité $\mathbf{V} \mathbf{V}^H = \mathbf{I}_n$.

Principe

Le principe de la méthode de Jacobi repose sur l'idée que toute matrice orthonormale $\mathbf{V} \in \mathbb{C}^{n \times n}$ peut s'écrire comme le produit de matrices de rotation de Givens complexes $\mathbf{G}(p, q, c, s)$,

définies pour $p < q$ par :

$$\mathbf{G}(p, q, c, s) = \begin{bmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \dots & c & \dots & \bar{s} & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & -s & \dots & c & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix} \begin{matrix} p \\ q \end{matrix} \quad (\text{A.4})$$

avec $(c, s) \in \mathbb{R} \times \mathbb{C}$ tels que $c^2 + |s|^2 = 1$.

La méthode de Jacobi consiste ainsi à appliquer successivement des rotations de Givens à la matrice \mathbf{A} jusqu'à minimisation du critère (A.3). A p et q fixés, une itération de la méthode consiste à :

$$\bullet \text{ minimiser}_{c,s} C_d(\mathbf{G}(p, q, c, s); \mathbf{A}) \quad (\text{A.5})$$

$$\bullet \mathbf{A} \leftarrow \mathbf{G}(p, q, c, s) \mathbf{A} \mathbf{G}(p, q, c, s)^H \quad (\text{A.6})$$

L'intérêt de la méthode réside dans le fait que l'étape de minimisation est possible de manière quasi-algébrique. Les résultats présentés dans les prochains paragraphes sont issus ou inspirés de [CS96, GL96].

Méthode

Observons l'effet d'une rotation de Givens sur la matrice \mathbf{A} . Soient p et q fixés avec $p < q$, on note :

$$\mathbf{B} = \mathbf{G}(p, q, c, s) \mathbf{A} \mathbf{G}(p, q, c, s)^H \quad (\text{A.7})$$

Un simple calcul montre que \mathbf{B} est partout égale à \mathbf{A} sauf sur les p^{eme} et q^{eme} lignes et colonnes :

$$\mathbf{B} = \begin{bmatrix} a_{ij} & c a_{ip} + s a_{iq} & a_{ij} & -\bar{s} a_{ip} + c a_{iq} & a_{ij} \\ c a_{pj} + \bar{s} a_{qj} & c^2 a_{pp} + |s|^2 a_{qq} & c a_{pj} + \bar{s} a_{qj} & c^2 a_{pq} - |s|^2 a_{qp} + c \bar{s} (a_{qq} - a_{pp}) & c a_{pj} + \bar{s} a_{qj} \\ a_{ij} & c a_{ip} + s a_{iq} & a_{ij} & -\bar{s} a_{ip} + c a_{iq} & a_{ij} \\ -s a_{pj} + c a_{qj} & c^2 a_{qp} - |s|^2 a_{pq} & -s a_{pj} + c a_{qj} & c^2 a_{qq} + |s|^2 a_{pp} - c s a_{pq} - c \bar{s} a_{qp} & -s a_{pj} + c a_{qj} \\ a_{ij} & c a_{ip} + s a_{iq} & a_{ij} & -\bar{s} a_{ip} + c a_{iq} & a_{ij} \end{bmatrix} \begin{matrix} p \\ q \end{matrix} \quad (\text{A.8})$$

Cherchons c et s qui minimisent $C_d(\mathbf{G}(p, q, c, s); \mathbf{A})$. On a :

$$C_d(\mathbf{G}(p, q, c, s); \mathbf{A}) = \text{off}(\mathbf{B}) \quad (\text{A.9})$$

$$= \sum_{1 \leq i \neq j \leq n} |b_{ij}|^2 \quad (\text{A.10})$$

$$= \|\mathbf{B}\|_F^2 - \sum_{i=1}^n |b_{ii}|^2 \quad (\text{A.11})$$

La norme de Frobenius étant invariante par rotation, on a $\|\mathbf{B}\|_F = \|\mathbf{A}\|_F$. En outre, les termes diagonaux de \mathbf{B} sont égaux aux termes diagonaux de \mathbf{A} sauf aux positions p et q . D'où :

$$\begin{aligned} \text{off}(\mathbf{B}) &= \|\mathbf{A}\|_F^2 - \left(\sum_{i=1}^n |a_{ii}|^2 - |a_{pp}|^2 - |a_{qq}|^2 + |b_{pp}|^2 + |b_{qq}|^2 \right) \\ &= \text{off}(\mathbf{A}) + |a_{pp}|^2 + |a_{qq}|^2 - |b_{pp}|^2 - |b_{qq}|^2 \end{aligned} \quad (\text{A.12})$$

La matrice \mathbf{A} ne dépendant pas de c et s , la minimisation suivant c et s de $C_d(\mathbf{G}(p, q, c, s); \mathbf{A})$ est donc ramenée à la maximisation de $|b_{pp}|^2 + |b_{qq}|^2$.

L'équation (A.12) peut-être développée d'avantage. En effet, grâce à l'identité triangulaire, on a :

$$|b_{pp}|^2 + |b_{qq}|^2 = \frac{1}{2} (|b_{pp} + b_{qq}|^2 + |b_{pp} - b_{qq}|^2) \quad (\text{A.13})$$

$$|a_{pp}|^2 + |a_{qq}|^2 = \frac{1}{2} (|a_{pp} + a_{qq}|^2 + |a_{pp} - a_{qq}|^2) \quad (\text{A.14})$$

En outre, la trace étant invariante par rotation, on a $b_{pp} + b_{qq} = a_{pp} + a_{qq}$. D'où :

$$\boxed{\text{off}(\mathbf{B}) = \text{off}(\mathbf{A}) + \frac{1}{2} [|a_{pp} - a_{qq}|^2 - |b_{pp} - b_{qq}|^2]} \quad (\text{A.15})$$

et donc :

$$\boxed{\text{minimiser}_{c,s} C_d(\mathbf{G}(p, q, c, s); \mathbf{A}) \iff \text{maximiser}_{c,s} C'_d(c, s) \stackrel{\text{def}}{=} |b_{pp} - b_{qq}|^2}$$

La dépendance de $C'_d(c, s)$ en p, q et \mathbf{A} a été omise afin d'alléger les notations par la suite.

Étudions à présent la maximisation de $C'_d(c, s)$. D'après l'équation (A.8), on a :

$$b_{pp} - b_{qq} = (c^2 - |s|^2)(a_{pp} - a_{qq}) + 2csa_{pq} + 2c\bar{s}a_{qp} \quad (\text{A.16})$$

En définissant :

$$\mathbf{v}(c, s) \stackrel{\text{def}}{=} [c^2 - |s|^2, cs + c\bar{s}, i(cs - c\bar{s})]^T \quad (\text{A.17})$$

$$\mathbf{h}(\mathbf{A}) \stackrel{\text{def}}{=} [a_{pp} - a_{qq}, a_{pq} + a_{qp}, i(a_{qp} - a_{pq})] \quad (\text{A.18})$$

on obtient :

$$b_{pp} - b_{qq} = \mathbf{h}(\mathbf{A}) \mathbf{v}(c, s) \quad (\text{A.19})$$

D'où :

$$C'_d(c, s) = \mathbf{v}(c, s)^T \mathbf{h}(\mathbf{A})^H \mathbf{h}(\mathbf{A}) \mathbf{v}(c, s) \quad (\text{A.20})$$

On remarquera que le vecteur $\mathbf{v}(c, s)$ est réel. De ce fait, et puisque $\mathbf{h}(\mathbf{A})^H \mathbf{h}(\mathbf{A})$ est par construction une matrice positive, il vient :

$$\boxed{C'_d(c, s) = \mathbf{v}(c, s)^T \text{Re} \{ \mathbf{h}(\mathbf{A})^H \mathbf{h}(\mathbf{A}) \} \mathbf{v}(c, s)} \quad (\text{A.21})$$

La maximisation de $C'_d(c, s)$ est donc ramenée à la maximisation de la forme quadratique positive $q_d(\mathbf{x}) = \mathbf{x}^T \mathbf{Q}_d \mathbf{x}$ sur le domaine réel $\{ \mathbf{v}(c, s) \mid (c, s) \in \mathbb{R} \times \mathbb{C}, c^2 + |s|^2 = 1 \}$, avec

$$\mathbf{Q}_d = \text{Re} \{ \mathbf{h}(\mathbf{A})^H \mathbf{h}(\mathbf{A}) \} \quad (\text{A.22})$$

Nous allons montrer que la maximisation de $q_d(\mathbf{x})$ sur le domaine précédent est équivalente à la maximisation de $q_d(\mathbf{x})$ sur la sphère unité, ce qui reviendra en fin de compte à calculer un vecteur propre associé à la plus grande valeur propre de \mathbf{Q}_d .

Propriété 9

Soit $\mathcal{E}_1 = \{\mathbf{v}(c, s) \mid (c, s) \in \mathbb{R} \times \mathbb{C}, c^2 + |s|^2 = 1\}$ et soit $\mathcal{E}_2 = \{[x, y, z]^T \mid (x, y, z) \in \mathbb{R}^3, x^2 + y^2 + z^2 = 1\}$. Alors on a $\mathcal{E}_1 = \mathcal{E}_2$.

Preuve de la propriété 9

- Soit $[u, v, w]^T \in \mathcal{E}_1$. Montrons que $[u, v, w]^T \in \mathcal{E}_2$.

Par définition il existe (c, s) tel que $u = c^2 - |s|^2, v = cs + c\bar{s}, w = i(cs - c\bar{s})$. On a :

$$u^2 + v^2 + w^2 = (c^2 - |s|^2)^2 + (cs + c\bar{s})^2 - (cs - c\bar{s})^2 \quad (\text{A.23})$$

$$= (c^2 - |s|^2)^2 + 4c^2 |s|^2 \quad (\text{A.24})$$

$$= (c^2 + |s|^2)^2 \quad (\text{A.25})$$

$$= 1 \quad (\text{A.26})$$

D'où $[u, v, w]^T \in \mathcal{E}_2$, et donc $\mathcal{E}_1 \subset \mathcal{E}_2$.

- Soit $[u, v, w]^T \in \mathcal{E}_2$. Montrons que $[u, v, w]^T \in \mathcal{E}_1$.

Si $u \neq -1$ on pose :

$$c = \sqrt{\frac{u+1}{2}}, \quad s = \frac{v - iw}{\sqrt{2(u+1)}} \quad (\text{A.27})$$

Si $u = -1$ (et donc $v = w = 0$) on pose :

$$c = 1, \quad s = 0 \quad (\text{A.28})$$

Un simple calcul montre alors que $c^2 + |s|^2 = 1$ et $[u, v, w]^T = [c^2 - |s|^2, cs + c\bar{s}, i(cs - c\bar{s})]^T = \mathbf{v}(c, s)$. D'où $[u, v, w]^T \in \mathcal{E}_1$, et donc $\mathcal{E}_2 \subset \mathcal{E}_1$.

• Finalement, $\mathcal{E}_1 \subset \mathcal{E}_2$ et $\mathcal{E}_2 \subset \mathcal{E}_1$, d'où $\mathcal{E}_2 = \mathcal{E}_1$. \square La maximisation de $C'_d(c, s)$ sur \mathcal{E}_1 est donc équivalente à la maximisation de $q_d(\mathbf{x})$ sur \mathcal{E}_2 . Or, la forme quadratique $q_d(\mathbf{x})$ est maximisée sur la sphère unité \mathcal{E}_2 par tout vecteur propre de norme unité associé à la plus grande valeur propre de \mathbf{Q}_d . Soit $[u_d^*, v_d^*, w_d^*]^T$ un tel vecteur propre, choisi tel que $u_d^* \geq 0$. D'après (A.27), les paramètres (c_d^*, s_d^*) qui minimisent le critère $C_d(\mathbf{G}(p, q, c, s); \mathbf{A})$ à p et q fixés sont donc :

$$\boxed{c_d^* = \sqrt{\frac{u_d^* + 1}{2}}, \quad s_d^* = \frac{v_d^* - iw_d^*}{\sqrt{2(u_d^* + 1)}}} \quad (\text{A.29})$$

Convergence de l'algorithme

Tel qu'il est construit, l'algorithme assure la décroissance du critère C_d à chaque itération. En effet, par définition de c_d^* et s_d^* , $\forall (c, s) \in \mathbb{R} \times \mathbb{C} \mid c^2 + |s|^2 = 1$ on a :

$$C_d(\mathbf{G}(p, q, c_d^*, s_d^*)) \leq C_d(\mathbf{G}(p, q, c, s)) \quad (\text{A.30})$$

En particulier, pour $(c, s) = (1, 0)$ on trouve :

$$\text{off}(\mathbf{B}) \leq \text{off}(\mathbf{A}) \quad (\text{A.31})$$

A chaque itération de l'algorithme, la matrice \mathbf{B} obtenue après rotations est donc «au moins aussi diagonale» que la matrice \mathbf{A} à la précédente itération. Puisque toute suite décroissante bornée dans \mathbb{R} converge, la convergence de l'algorithme est donc assurée. Cependant, ceci ne garantit pas que l'algorithme converge bien vers le minimum de C_d . La convergence vers le minimum a été observée en pratique mais nous ne sommes pas en mesure de la démontrer en théorie, dans le cas général. Cependant, nous allons voir que la convergence vers le minimum peut être démontrée pour des matrices hermitiennes.

Cas particulier : matrices hermitiennes

Si \mathbf{A} est une matrice hermitienne, alors elle est diagonalisable (et ses valeurs propres sont réelles). Dans ce cas, il est possible de donner l'expression explicite de $\text{off}(\mathbf{B}) - \text{off}(\mathbf{A})$ en fonction des seuls coefficients de \mathbf{A} . En effet, dans ce cas particulier, le vecteur $\mathbf{h}(\mathbf{A})$ est réel et donc $\mathbf{Q}_d = \mathbf{h}(\mathbf{A})^T \mathbf{h}(\mathbf{A})$. Nous avons vu précédemment que le maximum de $C'_d(c, s)$ est égal à la plus grande valeur propre de \mathbf{Q}_d . Or, \mathbf{Q}_d est par construction une matrice positive de rang 1, donc elle ne possède qu'une unique valeur propre non nulle que égale à sa trace. D'où :

$$|b_{pp}^* - b_{qq}^*|^2 = \text{trace}(\mathbf{Q}) = |a_{pp} - a_{qq}|^2 + |a_{pq} + a_{qp}|^2 + |a_{pq} - a_{qp}|^2 \quad (\text{A.32})$$

Donc, d'après (A.15) :

$$\begin{aligned} \text{off}(\mathbf{B}) - \text{off}(\mathbf{A}) &= \frac{1}{2} [|a_{pp} - a_{qq}|^2 - |b_{pp}^* - b_{qq}^*|^2] \\ &= -\frac{1}{2} [|a_{pq} + a_{qp}|^2 + |a_{pq} - a_{qp}|^2] \\ &= -(|a_{pq}|^2 + |a_{qp}|^2) \\ &= -2|a_{pq}|^2 \end{aligned}$$

Ainsi, dans le cas de matrices hermitiennes, la stricte décroissance du critère est assurée à chaque itération tant que l'un des coefficients non-diagonaux de \mathbf{A} est non-nul. Ceci assure donc la convergence de la méthode vers le minimum de C_d , qui vaut 0 dans le cas de matrices diagonalisables dans une base orthonormale.

Choix des rotations

Nous avons montré précédemment comment trouver à p et q fixés la rotation de Givens garantissant la décroissance maximale du critère C_d . Dans ce paragraphe nous décrivons succinctement deux choix de sélection de p et q .

Jacobi classique Cet algorithme prend son sens pour la diagonalisation d'une matrice hermitienne. Il consiste à choisir chaque itération p et q tel que $|a_{pq}|^2$ soit maximum afin d'assurer une décroissance maximale de C_d , voir l'algorithme 1.

Algorithme 1 Jacobi classique

```

 $\mathbf{U} = \mathbf{I}_n$ 
 $\epsilon = \text{tol} \cdot \|\mathbf{A}\|_F$ 
tant que  $\text{off}(\mathbf{A}) > \epsilon$  faire
  Choisir  $p$  et  $q$  tels que  $|a_{pq}|$  est maximal
  Calculer  $c_d^*$  et  $s_d^*$ 
   $\mathbf{A} \leftarrow \mathbf{G}(p, q, c_d^*, s_d^*) \mathbf{A} \mathbf{G}(p, q, c_d^*, s_d^*)^H$ 
   $\mathbf{U} \leftarrow \mathbf{U} \mathbf{G}(p, q, c_d^*, s_d^*)$ 
fin tant que

```

Jacobi cyclique L'algorithme Jacobi cyclique consiste à balayer systématiquement les lignes de la matrice \mathbf{A} . Pour des matrices hermitiennes il est possible de montrer qu'il est plus rapide que l'algorithme de Jacobi classique bien qu'il nécessite de rechercher le terme $|a_{pq}|^2$ maximum à chaque itération (voir [GL96]), voir l'algorithme 2. Le parcours complet de toutes les positions (p, q) non-diagonales est appelé un *balayage*.

Algorithme 2 Jacobi cyclique

```

U = In
 $\epsilon = \text{tol} \cdot \|\mathbf{A}\|_F$ 
tant que  $\text{off}(\mathbf{A}) > \epsilon$  faire
  pour  $p = 1 : n - 1$  faire
    pour  $q = p + 1 : n$  faire
      Calculer  $c_d^*$  et  $s_d^*$ 
       $\mathbf{A} \leftarrow \mathbf{G}(p, q, c_d^*, s_d^*) \mathbf{A} \mathbf{G}(p, q, c_d^*, s_d^*)^H$ 
       $\mathbf{U} \leftarrow \mathbf{U} \mathbf{G}(p, q, c_d^*, s_d^*)$ 
    fin pour
  fin pour
fin tant que

```

A.1.2 Base orthonormale réelle

Nous présentons dans ce paragraphe quelques simplifications de la méthode précédente dans le cas particulier où l'on cherche à diagonaliser \mathbf{A} dans une base orthonormale réelle, c-à-d, $\mathbf{U} \in \mathbb{R}^{n \times n}$ et $\mathbf{D} \in \mathbb{C}^{n \times n}$. Dans ce cas, on peut rechercher une telle base sous la forme d'un produit de matrices de rotation de Givens *réelles* $\mathbf{G}_r(p, q, c, s)$, définies pour $p < q$ par :

$$\mathbf{G}_r(p, q, c, s) = \begin{bmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \dots & c & \dots & s & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & -s & \dots & c & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix} \begin{matrix} p \\ q \end{matrix} \quad (\text{A.33})$$

avec $(c, s) \in \mathbb{R}^2$ tels que $c^2 + s^2 = 1$. On procède toujours par itérations, et à p et q fixés, une itération consiste à :

$$\bullet \quad \underset{c, s}{\text{minimiser}} \quad C_d(\mathbf{G}_r(p, q, c, s); \mathbf{A}) \quad (\text{A.34})$$

$$\bullet \quad \mathbf{A} \leftarrow \mathbf{G}_r(p, q, c, s) \mathbf{A} \mathbf{G}_r(p, q, c, s)^T \quad (\text{A.35})$$

Comme précédemment, minimiser $C_d(\mathbf{G}_r(p, q, c, s); \mathbf{A})$ revient à maximiser $C'_d(c, s) = |b_{pp} - b_{qq}|^2$. Mais d'après l'équation (A.16) dans le cas réel on a :

$$b_{pp} - b_{qq} = (c^2 - s^2)(a_{pp} - a_{qq}) + 2cs(a_{pq} + a_{qp}) \quad (\text{A.36})$$

En définissant :

$$\mathbf{v}_r(c, s) = [c^2 - s^2, 2cs]^T \quad (\text{A.37})$$

$$\mathbf{h}_r(\mathbf{A}) = [a_{pp} - a_{qq}, a_{pq} + a_{qp}] \quad (\text{A.38})$$

on obtient :

$$b_{pp} - b_{qq} = \mathbf{h}_r(\mathbf{A}) \mathbf{v}_r(c, s) \quad (\text{A.39})$$

D'où :

$$C'_d(c, s) = \mathbf{v}_r(c, s)^T \mathbf{h}_r(\mathbf{A})^H \mathbf{h}_r(\mathbf{A}) \mathbf{v}_r(c, s) \quad (\text{A.40})$$

$$= \mathbf{v}_r(c, s)^T \text{Re} \{ \mathbf{h}_r(\mathbf{A})^H \mathbf{h}_r(\mathbf{A}) \} \mathbf{v}_r(c, s) \quad (\text{A.41})$$

De la même façon que précédemment, la maximisation de $C'_d(c, s)$ est équivalente à la maximisation de la forme quadratique $q_r(\mathbf{x}) = \mathbf{x}^T \mathbf{Q}_{dr} \mathbf{x}$ sur la sphère unité, avec

$$\mathbf{Q}_{dr} = \text{Re} \{ \mathbf{h}_r(\mathbf{A})^T \mathbf{h}_r(\mathbf{A}) \} \quad (\text{A.42})$$

Soit $[u_{dr}^*, v_{dr}^*]^T$ un vecteur propre de \mathbf{Q}_{dr} associé à la plus grande valeur propre, choisi tel que $u_{dr}^* \geq 0$. D'après l'équation (A.27), les paramètres (c_{dr}^*, s_{dr}^*) qui minimisent le critère $C_d(\mathbf{G}_r(p, q, c, s); \mathbf{A})$ à p et q fixés sont alors :

$$\boxed{c_{dr}^* = \sqrt{\frac{u_{dr}^* + 1}{2}}, \quad s_{dr}^* = \frac{v_{dr}^*}{\sqrt{2(u_{dr}^* + 1)}}} \quad (\text{A.43})$$

Remarquons qu'il est possible de donner les expressions algébriques de c_{dr}^* et s_{dr}^* en utilisant les expressions algébriques des éléments propres d'une matrice 2×2 .

A.2 Diagonalisation simultanée approchée

A.2.1 Base orthonormale complexe

On considère ici le problème de la *diagonalisation simultanée approchée* (appelée encore *diagonalisation conjointe approchée*) d'un ensemble de matrices. Soit $\mathcal{A} = \{\mathbf{A}_1, \dots, \mathbf{A}_K\}$ un ensemble de K matrices $\in \mathbb{C}^{n \times n}$. Le problème consiste à trouver une matrice orthonormale $\mathbf{U} \in \mathbb{C}^{n \times n}$ telle que $\forall k = 1, \dots, K$, les matrices :

$$\mathbf{U} \mathbf{A}_k \mathbf{U}^H = \mathbf{D}_k \quad (\text{A.44})$$

soient les plus diagonales possible au sens du critère (A.2). En d'autres termes, on cherche à minimiser le critère :

$$C_{ds}(\mathbf{V}; \mathcal{A}) = \sum_{k=1}^K \text{off}(\mathbf{V} \mathbf{A}_k \mathbf{V}^H) \quad (\text{A.45})$$

où $\mathbf{V} \in \mathbb{C}^{n \times n}$, sous la contrainte d'orthonormalité $\mathbf{V} \mathbf{V}^H = \mathbf{I}_n$.

Principe

Nous décrivons ici l'algorithme de type Jacobi présenté dans [CS96]. Il est une généralisation simple et efficace de la méthode de Jacobi pour la diagonalisation approchée d'une matrice présentée au paragraphe A.1.1. La base orthonormale commune \mathbf{U} est estimée itérativement par produit de rotations de Givens complexes qui font décroître le critère $C_{ds}(\mathbf{V}; \mathcal{A})$. A p et q fixés, une itération de la méthode consiste à :

$$\bullet \quad \underset{c, s}{\text{minimiser}} \quad C_{ds}(\mathbf{G}(p, q, c, s); \mathcal{A}) \quad (\text{A.46})$$

$$\bullet \quad \forall k, \mathbf{A}_k \leftarrow \mathbf{G}(p, q, c, s) \mathbf{A}_k \mathbf{G}(p, q, c, s)^H \quad (\text{A.47})$$

Méthode

$\forall k \in \llbracket 1, K \rrbracket$, on note :

$$\mathbf{B}_k = \mathbf{G}(p, q, c, s) \mathbf{A}_k \mathbf{G}(p, q, c, s)^H \quad (\text{A.48})$$

On a donc :

$$C_{ds}(\mathbf{G}(p, q, c, s); \mathcal{A}) = \sum_{k=1}^K \text{off}(\mathbf{B}_k) \quad (\text{A.49})$$

En utilisant la notation $\mathbf{A}_k = \{a_{kij}\}$, d'après (A.12) on a :

$$C_{ds}(\mathbf{G}(p, q, c, s); \mathcal{A}) = \sum_{k=1}^K \text{off}(\mathbf{A}_k) + |a_{kpp}|^2 + |a_{kqq}|^2 - |b_{kpp}|^2 - |b_{kqq}|^2 \quad (\text{A.50})$$

$$= \sum_{k=1}^K \text{off}(\mathbf{A}_k) + \frac{1}{2} [|a_{kpp} - a_{kqq}|^2 - |b_{kpp} - b_{kqq}|^2] \quad (\text{A.51})$$

De la même façon qu'au paragraphe A.1.1, la minimisation de $C_{ds}(\mathbf{G}(p, q, c, s); \mathcal{A})$ est ramenée à la maximisation de $C'_{ds}(c, s) \stackrel{\text{def}}{=} \sum_{k=1}^K |b_{kpp} - b_{kqq}|^2$. Avec les mêmes notations que précédemment, on a alors :

$$\begin{aligned} C'_{ds}(c, s) &= \sum_{k=1}^K \mathbf{v}(c, s)^T \mathbf{h}(\mathbf{A}_k)^H \mathbf{h}(\mathbf{A}_k) \mathbf{v}(c, s) \\ &= \mathbf{v}(c, s)^T \left[\sum_{k=1}^K \mathbf{h}(\mathbf{A}_k)^H \mathbf{h}(\mathbf{A}_k) \right] \mathbf{v}(c, s) \\ &= \mathbf{v}(c, s)^T \text{Re} \left\{ \sum_{k=1}^K \mathbf{h}(\mathbf{A}_k)^H \mathbf{h}(\mathbf{A}_k) \right\} \mathbf{v}(c, s) \end{aligned}$$

La maximisation de $C'_{ds}(c, s)$ sur \mathcal{E}_1 est donc équivalente à la maximisation de la forme quadratique $q_{ds}(\mathbf{x}) = \mathbf{x}^T \mathbf{Q}_{ds} \mathbf{x}$ sur \mathcal{E}_2 , avec

$$\mathbf{Q}_{ds} = \text{Re} \left\{ \sum_{k=1}^K \mathbf{h}(\mathbf{A}_k)^H \mathbf{h}(\mathbf{A}_k) \right\} \quad (\text{A.52})$$

Soit $[u_{ds}^*, v_{ds}^*, w_{ds}^*]^T$ un vecteur propre de \mathbf{Q}_{ds} associé à la plus grande valeur propre, choisi tel que $u_{ds}^* \geq 0$. D'après (A.27), les paramètres (c_{ds}^*, s_{ds}^*) qui minimisent le critère $C_{ds}(\mathbf{G}(p, q, c, s); \mathcal{A})$ à p et q fixés sont :

$$c_{ds}^* = \sqrt{\frac{u_{ds}^* + 1}{2}}, \quad s_{ds}^* = \frac{v_{ds}^* - i w_{ds}^*}{\sqrt{2(u_{ds}^* + 1)}} \quad (\text{A.53})$$

Convergence de l'algorithme

Comme pour la diagonalisation approchée d'une matrice (paragraphe A.1.1), l'algorithme tel qu'il est construit garantit la décroissance du critère C_{ds} à chaque itération. En effet, $\forall (c, s) \in \mathbb{R} \times \mathbb{C} \mid c^2 + |s|^2 = 1$, on a :

$$C_{ds}(\mathbf{G}(p, q, c_{ds}^*, s_{ds}^*)) \leq C_{ds}(\mathbf{G}(p, q, c, s)) \quad (\text{A.54})$$

En particulier, pour $(c, s) = (1, 0)$ on trouve :

$$\sum_{k=1}^K \text{off}(\mathbf{B}_k) \leq \sum_{k=1}^K \text{off}(\mathbf{A}_k) \quad (\text{A.55})$$

La convergence de l'algorithme est donc garantie, mais, comme pour la diagonalisation d'une simple matrice, rien n'assure la convergence vers le minimum de C_{ds} . La convergence théorique de l'algorithme peut être facilement démontrée pour un ensemble de matrices hermitiennes possédant une base orthonormale commune, mais nous ne connaissons pas de démonstration théorique de la convergence vers le minimum dans le cas général. Toutefois, avec le choix des rotations cyclique, la convergence vers le minimum a toujours été observée en pratique.

Une implémentation MATLAB de l'algorithme présenté est disponible sur la page Internet personnelle de J.-F. Cardoso [Car]. Dans cette version, le choix des rotations est cyclique et le critère d'arrêt porte sur s_{ds}^* : l'algorithme s'arrête lorsque toutes les valeurs de s_{ds}^* sont inférieures à un seuil (par défaut 10^{-8}) durant un balayage.

A.2.2 Base orthonormale réelle

Nous présentons dans ce paragraphe les simplifications possible de la méthode lorsque l'on cherche à diagonaliser simultanément l'ensemble \mathcal{A} dans une base orthonormale commune *réelle* (c-à-d, $\mathbf{U} \in \mathbb{R}^{n \times n}$ et $\forall k \in \llbracket 1, K \rrbracket$, $\mathbf{D}_k \in \mathbb{C}^{n \times n}$). Dans ce cas, comme dans la partie A.1.2, le critère à maximiser est :

$$\begin{aligned} C'_{ds}(p, q) &= \sum_{k=1}^K \mathbf{v}_r(c, s)^T \mathbf{h}_r(\mathbf{A}_k)^T \mathbf{h}_r(\mathbf{A}_k) \mathbf{v}_r(c, s) \\ &= \mathbf{v}_r(c, s)^T \left[\sum_{k=1}^K \mathbf{h}_r(\mathbf{A}_k)^T \mathbf{h}_r(\mathbf{A}_k) \right] \mathbf{v}_r(c, s) \end{aligned}$$

La maximisation de ce critère est donc équivalente à la maximisation de la forme quadratique $q(\mathbf{x}) = \mathbf{x}^T \mathbf{Q}_{dsr} \mathbf{x}$, avec

$$\mathbf{Q}_{dsr} = \text{Re} \left\{ \sum_{k=1}^K \mathbf{h}_r(\mathbf{A}_k)^T \mathbf{h}_r(\mathbf{A}_k) \right\}$$

Soit $[u_{dsr}^*, v_{dsr}^*]^T$ un vecteur propre de \mathbf{Q}_{dsr} associé à la plus grande valeur propre de \mathbf{Q}_{dsr} , choisi tel que $u_{dsr}^* \geq 0$. D'après (A.27), les paramètres (c_{dsr}^*, s_{dsr}^*) qui minimisent le critère $C_{ds}(\mathbf{G}_r(p, q, c, s); \mathcal{A})$ à p et q fixés sont alors :

$$\boxed{c_{dsr}^* = \sqrt{\frac{u_{dsr}^* + 1}{2}}, \quad s_{dsr}^* = \frac{v_{dsr}^*}{\sqrt{2(u_{dsr}^* + 1)}}} \quad (\text{A.56})$$

A.3 Bloc-diagonalisation approchée

Nous proposons dans les prochains paragraphes deux méthodes de Jacobi pour la *bloc*-diagonalisation approchée et la *bloc*-diagonalisation simultanée approchée de matrices. Ces méthodes sont dans la lignée des méthodes de diagonalisation exposées précédemment. Cependant nous avons rencontré dans le cas *bloc* des problèmes de convergence qui n'ont pas encore été élucidés et qui seront discutés à la fin de cette partie.

A.3.1 Base orthonormale complexe

Nous présentons une méthode de type Jacobi pour la bloc-diagonalisation approchée d'une matrice $\mathbf{A} \in \mathbb{C}^{n \times n}$ dans une base orthonormale. Le problème consiste à trouver une matrice orthonormale $\mathbf{U} \in \mathbb{C}^{n \times n}$ telle que

$$\mathbf{U} \mathbf{A} \mathbf{U}^H = \mathbf{D} \quad (\text{A.57})$$

soit *le plus bloc-diagonal possible*. On notera L la dimension (fixée) des blocs de \mathbf{D} et $m = n/L$ le nombre de blocs.

Notons :

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \dots & \mathbf{A}_{1m} \\ \vdots & & \vdots \\ \mathbf{A}_{m1} & \dots & \mathbf{A}_{mm} \end{bmatrix} \quad (\text{A.58})$$

où $\forall (i, j) \in \llbracket 1, m \rrbracket^2$, \mathbf{A}_{ij} est une matrice de dimensions $L \times L$. On définit le critère de "bloc-diagonalité" suivant :

$$\text{boff}(\mathbf{A}) \stackrel{\text{def}}{=} \sum_{1 \leq i \neq j \leq m} \|\mathbf{A}_{ij}\|_F^2 \quad (\text{A.59})$$

On cherche la matrice \mathbf{U} par minimisation du critère suivant :

$$C_{bd}(\mathbf{V}; \mathbf{A}) = \text{boff}(\mathbf{V} \mathbf{A} \mathbf{V}^H) \quad (\text{A.60})$$

sous la contrainte d'orthonormalité $\mathbf{V} \mathbf{V}^H = \mathbf{I}_n$. Comme précédemment, le principe de la méthode est d'appliquer successivement des rotations de Givens à la matrice \mathbf{A} jusqu'à minimisation du critère (A.60). A p et q fixés, une itération de la méthode consiste à :

$$\bullet \quad \underset{c,s}{\text{minimiser}} \quad C_{bd}(\mathbf{G}(p, q, c, s); \mathbf{A}) \quad (\text{A.61})$$

$$\bullet \quad \mathbf{A} \leftarrow \mathbf{G}(p, q, c, s) \mathbf{A} \mathbf{G}(p, q, c, s)^H \quad (\text{A.62})$$

Notations

Soient I_1, \dots, I_m m sous-intervalles consécutifs de $\llbracket 1, n \rrbracket$ définis par :

$$I_1 = \llbracket 1, L \rrbracket \quad (\text{A.63})$$

$$\vdots \quad (\text{A.64})$$

$$I_i = \llbracket (i-1)L + 1, iL \rrbracket \quad (\text{A.65})$$

$$\vdots \quad (\text{A.66})$$

$$I_m = \llbracket (m-1)L + 1, mL \rrbracket \quad (\text{A.67})$$

Soit $i(k)$ la fonction définie par :

$$i : \llbracket 1, n \rrbracket \rightarrow \llbracket 1, m \rrbracket \quad (\text{A.68})$$

$$k \rightarrow \left\lceil \frac{k}{L} \right\rceil \quad (\text{A.69})$$

où $\lceil x \rceil$ définit la partie entière supérieure de x (ex : $\lceil 2 \rceil = 2$, $\lceil 2.1 \rceil = 3$). La fonction $i(k)$ renvoie l'indice i de l'intervalle I_i auquel k appartient.

Méthode

Soient $(p, q) \in \llbracket 1, n \rrbracket^2$, $p < q$, on pose :

$$\mathbf{B} = \mathbf{G}(p, q, c, s) \mathbf{A} \mathbf{G}(p, q, c, s)^H \quad (\text{A.70})$$

D'après (A.59), (A.60) et avec $\|\mathbf{B}\|_F^2 = \sum_{ij} \|\mathbf{B}_{ij}\|_F^2$, on a :

$$C_{bd}(\mathbf{G}(p, q, c, s); \mathbf{A}) = \text{boff}(\mathbf{B}) \quad (\text{A.71})$$

$$= \|\mathbf{B}\|_F^2 - \sum_{i=1}^m \|\mathbf{B}_{ii}\|_F^2 \quad (\text{A.72})$$

$$= \|\mathbf{A}\|_F^2 - \sum_{i=1}^m \|\mathbf{B}_{ii}\|_F^2 \quad (\text{A.73})$$

$$(\text{A.74})$$

Supposons que p et q appartiennent au même sous-intervalle I_i , c-à-d, $i(p) = i(q)$. D'après l'équation (A.8), \mathbf{B} est partout égale à \mathbf{A} sauf sur les p^{eme} et q^{eme} lignes et colonnes. D'où :

$$\sum_{i=1}^m \|\mathbf{B}_{ii}\|_F^2 = \sum_{i=1}^m \|\mathbf{A}_{ii}\|_F^2 - \|\mathbf{A}_{i(p)i(p)}\|_F^2 + \|\mathbf{B}_{i(p)i(p)}\|_F^2 \quad (\text{A.75})$$

Or, puisque p et q appartiennent au même intervalle, $\|\mathbf{B}_{i(p)i(p)}\|_F = \|\mathbf{A}_{i(p)i(p)}\|_F$ (invariance de la norme de Frobenius par rotation).

Donc, si $i(p) = i(q)$, $\sum_{i=1}^m \|\mathbf{B}_{ii}\|_F^2 = \sum_{i=1}^m \|\mathbf{A}_{ii}\|_F^2$, c-à-d, $\text{boff}(\mathbf{B}) = \text{boff}(\mathbf{A})$, et donc finalement C_{bd} reste constant. Par la suite on supposera donc $i(p) \neq i(q)$, c-à-d, $q - p \geq L$.

Dans ce cas, on a alors :

$$C_{bd}(\mathbf{G}(p, q, c, s); \mathbf{A}) = \quad (\text{A.76})$$

$$\begin{aligned} & \left\| \mathbf{A} \right\|_F^2 - \left\{ \sum_{k=1}^m \|\mathbf{A}_{kk}\|_F^2 - \|\mathbf{A}_{i(p)i(p)}\|_F^2 - \|\mathbf{A}_{i(q)i(q)}\|_F^2 + \|\mathbf{B}_{i(p)i(p)}\|_F^2 + \|\mathbf{B}_{i(q)i(q)}\|_F^2 \right\} \\ &= \text{boff}(\mathbf{A}) + \|\mathbf{A}_{i(p)i(p)}\|_F^2 + \|\mathbf{A}_{i(q)i(q)}\|_F^2 - \|\mathbf{B}_{i(p)i(p)}\|_F^2 - \|\mathbf{B}_{i(q)i(q)}\|_F^2 \end{aligned} \quad (\text{A.77})$$

La minimisation de $C_{bd}(\mathbf{G}(p, q, c, s); \mathbf{A})$ est donc équivalente à la maximisation de $\|\mathbf{B}_{i(p)i(p)}\|_F^2 + \|\mathbf{B}_{i(q)i(q)}\|_F^2$. Mais puisque seules les p^{eme} et q^{eme} lignes et colonnes de \mathbf{B} dépendent de c et s , la minimisation de C_{bd} est finalement ramenée à la maximisation du critère :

$$C'_{bd}(c, s) \stackrel{\text{def}}{=} |b_{pp}|^2 + |b_{qq}|^2 + \sum_{j \in I_{i(p)}, j \neq p} |b_{pj}|^2 + |b_{jp}|^2 + \sum_{j \in I_{i(q)}, j \neq q} |b_{qj}|^2 + |b_{jq}|^2 \quad (\text{A.78})$$

où l'on rappelle les expressions de b_{pp} , b_{qq} , b_{pj} , b_{jp} , b_{qj} et b_{jq} (d'après (A.8)) :

$$b_{pp} = c^2 a_{pp} + |s|^2 a_{qq} + c s a_{pq} + c \bar{s} a_{qp} \quad (\text{A.79})$$

$$b_{qq} = c^2 a_{qq} + |s|^2 a_{pp} - c s a_{pq} - c \bar{s} a_{qp} \quad (\text{A.80})$$

$$b_{pj} = c a_{pj} + \bar{s} a_{qj} \quad (j \neq p) \quad (\text{A.81})$$

$$b_{jp} = c a_{jp} + s a_{jq} \quad (j \neq p) \quad (\text{A.82})$$

$$b_{qj} = -s a_{pj} + c a_{qj} \quad (j \neq q) \quad (\text{A.83})$$

$$b_{jq} = -\bar{s} a_{jp} + c a_{jq} \quad (j \neq q) \quad (\text{A.84})$$

Comme précédemment, la dépendance de $C'_{bd}(c, s)$ en p , q et \mathbf{A} a été omise afin d'alléger les notations.

Dans [BAMH98] il est montré que la maximisation de $C'_{bd}(c, s)$ peut être ramenée à la maximisation contrainte d'une forme quadratique. Cette maximisation contrainte est réalisée au moyen de multiplicateurs de Lagrange. Le calcul de ces derniers nécessite le calcul des racines d'un polynôme d'ordre 6 dans le cas complexe (c-à-d, $\mathbf{U} \in \mathbb{C}^{n \times n}$), et d'ordre 4 dans le cas réel (c-à-d, $\mathbf{U} \in \mathbb{R}^{n \times n}$). Pour le cas complexe, une autre alternative, basée sur des approximations tensorielle de rang 1, est proposée dans [LFMV02].

Nous décrivons par la suite une méthode de bloc-diagonalisation approchée de matrices réelles dans une base orthonormale réelle. L'approche nous présentons, proposée dans [LFMV02], se résume à chaque étape de la méthode au calcul des racines d'un polynôme d'ordre 4 et possède l'avantage d'être plus directe que la méthode proposée dans [BAMH98].

A.3.2 Matrices réelles

Soit $\mathbf{A} \in \mathbb{R}^{n \times n}$ que l'on cherche à bloc-diagonaliser dans une base orthonormale réelle (c-à-d, $\mathbf{U} \in \mathbb{R}^{n \times n}$). Comme pour la diagonalisation, on recherche une telle base sous la forme d'un produit de matrices de Givens réelles $\mathbf{G}_r(p, q, c, s)$. $C'_{bd}(c, s)$ est alors égal à $C'_{bdr}(c, s)$, défini par :

$$C'_{bdr}(c, s) \stackrel{\text{def}}{=} (b_{pp})^2 + (b_{qq})^2 + \sum_{j \in I_{i(p)}, j \neq p} (b_{pj})^2 + (b_{jp})^2 + \sum_{j \in I_{i(q)}, j \neq q} (b_{qj})^2 + (b_{jq})^2 \quad (\text{A.85})$$

avec :

$$(b_{pp})^2 = (c^2 a_{pp} + s^2 a_{qq} + c s (a_{pq} + a_{qp}))^2 \quad (\text{A.86})$$

$$(b_{qq})^2 = (c^2 a_{qq} + s^2 a_{pp} - c s (a_{pq} + a_{qp}))^2 \quad (\text{A.87})$$

$$(b_{pj})^2 = (c a_{pj} + s a_{qj})^2 (c^2 + s^2) \quad (j \neq p) \quad (\text{A.88})$$

$$(b_{jp})^2 = (c a_{jp} + s a_{jq})^2 (c^2 + s^2) \quad (j \neq p) \quad (\text{A.89})$$

$$(b_{qj})^2 = (-s a_{pj} + c a_{qj})^2 (c^2 + s^2) \quad (j \neq q) \quad (\text{A.90})$$

$$(b_{jq})^2 = (-s a_{jp} + c a_{jq})^2 (c^2 + s^2) \quad (j \neq q) \quad (\text{A.91})$$

$C'_{bdr}(c, s)$ est un polynôme d'ordre 4 en c et s . Les expressions de $(b_{pj})^2, (b_{jp})^2, (b_{qj})^2$ et $(b_{jq})^2$ ont été multipliées par $(c^2 + s^2) = 1$ afin de rendre $C'_{bdr}(c, s)$ homogène en c et s (c-à-d que tous les termes de $C'_{bdr}(c, s)$ sont de la forme $c^i s^j$ avec $i + j = 4$).

Pour tout $(c, s) \in \mathbb{R}^2$ tel que $c^2 + s^2 = 1$, il existe $\theta \in [0, 2\pi[$ tel que $c = \cos \theta$ et $s = \sin \theta$. $C'_{bdr}(c, s)$ peut donc s'exprimer en fonction du seul paramètre θ . En développant l'expression (A.85) on obtient alors l'expression suivante, à maximiser en θ :

$$C'_{bdr}(\theta) = q_{40} \cos^4 \theta + q_{04} \sin^4 \theta + q_{31} \cos^3 \theta \sin \theta + q_{13} \cos \theta \sin^3 \theta + q_{22} \cos^2 \theta \sin^2 \theta \quad (\text{A.92})$$

avec :

$$q_{40} =$$

$$a_{pp}^2 + a_{qq}^2 + \sum_{j \in I_{i(p)}, j \neq p} a_{pj}^2 + a_{jp}^2 + \sum_{j \in I_{i(q)}, j \neq q} a_{qj}^2 + a_{jq}^2 \quad (\text{A.93})$$

$$q_{04} =$$

$$a_{pp}^2 + a_{qq}^2 + \sum_{j \in I_{i(p)}, j \neq p} a_{qj}^2 + a_{jq}^2 + \sum_{j \in I_{i(q)}, j \neq q} a_{pj}^2 + a_{jp}^2 \quad (\text{A.94})$$

$$q_{31} =$$

$$2 \left((a_{pp} - a_{qq}) (a_{pq} + a_{qp}) + \sum_{j \in I_{i(p)}, j \neq p} a_{pj} a_{qj} + a_{jp} a_{jq} - \sum_{j \in I_{i(q)}, j \neq q} a_{pj} a_{qj} + a_{jp} a_{jq} \right) \quad (\text{A.95})$$

$$q_{13} =$$

$$2 \left((a_{qq} - a_{pp}) (a_{pq} + a_{qp}) + \sum_{j \in I_{i(p)}, j \neq p} a_{pj} a_{qj} + a_{jp} a_{jq} - \sum_{j \in I_{i(q)}, j \neq q} a_{pj} a_{qj} + a_{jp} a_{jq} \right) \quad (\text{A.96})$$

$$q_{22} =$$

$$2(a_{pq} + a_{qp})^2 + 4a_{pp} a_{qq} + \sum_{j \in I_{i(p)}, j \neq p} (a_{pj}^2 + a_{qj}^2) + (a_{jp}^2 + a_{jq}^2) + \sum_{j \in I_{i(q)}, j \neq q} (a_{pj}^2 + a_{qj}^2) + (a_{jp}^2 + a_{jq}^2) \quad (\text{A.97})$$

La fonction $C'_{bdr}(\theta)$ est périodique de période π . On en cherchera donc le maximum sur l'intervalle $]-\frac{\pi}{2}, \frac{\pi}{2}]$, c-à-d, pour $\cos \theta \geq 0$. La dérivée de $C'_{bd}(\theta)$ s'écrit :

$$\begin{aligned} \frac{\delta C'_{bd}}{\delta \theta}(\theta) = & q_{31} \cos^4 \theta - q_{13} \sin^4 \theta - 2(2q_{40} - q_{22}) \cos^3 \theta \sin \theta + 2(2q_{04} - q_{22}) \cos \theta \sin^3 \theta \\ & + 3(q_{13} - q_{31}) \cos^2 \theta \sin^2 \theta \quad (\text{A.98}) \end{aligned}$$

En divisant l'expression de $C'_{bdr}(\theta)$ par $\cos^4 \theta$ pour $\theta \neq \frac{\pi}{2}$, il vient :

$$\frac{\delta C'_{bdr}}{\delta \theta}(\theta) = 0 \iff \begin{cases} \theta = \frac{\pi}{2} & \text{si } q_{13} = 0 \\ P(\tan \theta) = 0 & \text{si } \theta \in]-\frac{\pi}{2}, \frac{\pi}{2}[\end{cases} \quad (\text{A.99})$$

où $P(x)$ est le polynôme défini par :

$$P(x) = q_{13} x^4 - 2(2q_{04} - q_{22}) x^3 - 3(q_{13} - q_{31}) x^2 + 2(2q_{40} - q_{22}) x - q_{31} \quad (\text{A.100})$$

La recherche du maximum de $C'_{bdr}(\theta)$ est donc ramenée à la recherche des racines d'un polynôme d'ordre 4. Les racines de P peuvent être estimées numériquement et on retient alors la racine réelle θ_{bdr}^* dont la valeur de l'arctangente maximise $C'_{bdr}(\theta)$.

A.4 Bloc-diagonalisation simultanée approchée

A.4.1 Base orthonormale complexe

Principe

On considère maintenant le problème de la *bloc-diagonalisation simultanée* approchée d'un ensemble de K matrices $\in \mathbb{C}^{n \times n}$ $\mathcal{A} = \{\mathbf{A}_1, \dots, \mathbf{A}_K\}$. Le problème consiste à trouver une matrice orthonormale $\mathbf{U} \in \mathbb{C}^{n \times n}$ telle que $\forall k \in \llbracket 1, K \rrbracket$, les matrices :

$$\mathbf{U} \mathbf{A}_k \mathbf{U}^H = \mathbf{B}_k \quad (\text{A.101})$$

soient les plus bloc-diagonales possible au sens du critère (A.59).

La méthode que nous présentons découle naturellement des méthodes de diagonalisation simultanée et de bloc-diagonalisation proposées aux paragraphes A.2 et A.3. En utilisant comme précédemment les notations $\mathbf{A}_k = \{a_{kij}\}$ et, pour $k \in \llbracket 1, K \rrbracket$:

$$\mathbf{A}_k = \begin{bmatrix} \mathbf{A}_{k11} & \dots & \mathbf{A}_{k1m} \\ \vdots & & \vdots \\ \mathbf{A}_{km1} & \dots & \mathbf{A}_{kmm} \end{bmatrix} \quad (\text{A.102})$$

où $\forall (i, j) \in \llbracket 1, m \rrbracket^2$, \mathbf{A}_{kij} est une matrice de dimensions $L \times L$, la matrice \mathbf{U} peut être estimée par minimisation du critère :

$$C_{bds}(\mathbf{V}; \mathcal{A}) = \sum_{i=1}^K \text{boff}(\mathbf{V} \mathbf{A}_i \mathbf{V}^H) \quad (\text{A.103})$$

sous la contrainte $\mathbf{V} \mathbf{V}^H = \mathbf{I}_n$. A p et q fixés, une itération de la méthode consiste à :

$$\bullet \text{ minimiser}_{c,s} C_{bds}(\mathbf{G}(p, q, c, s); \mathcal{A}) \quad (\text{A.104})$$

$$\bullet \forall k, \mathbf{A}_k \leftarrow \mathbf{G}(p, q, c, s) \mathbf{A}_k \mathbf{G}(p, q, c, s)^H \quad (\text{A.105})$$

Méthode

$\forall k = 1, \dots, K$, on note :

$$\mathbf{B}_k = \mathbf{G}(p, q, c, s) \mathbf{A}_k \mathbf{G}(p, q, c, s)^H \quad (\text{A.106})$$

On a donc :

$$C_{bds}(\mathbf{G}(p, q, c, s); \mathcal{A}) = \sum_{k=1}^K \text{boff}(\mathbf{B}_k) \quad (\text{A.107})$$

D'après (A.77), pour p et q tels que $i(p) \neq i(q)$, on alors :

$$C_{bds}(\mathbf{G}(p, q, c, s); \mathcal{A}) = \sum_{k=1}^K \text{boff}(\mathbf{A}_k) + \|\mathbf{A}_{ki(p)i(p)}\|_F^2 + \|\mathbf{A}_{ki(q)i(q)}\|_F^2 - \|\mathbf{B}_{ki(p)i(p)}\|_F^2 - \|\mathbf{B}_{ki(q)i(q)}\|_F^2 \quad (\text{A.108})$$

La minimisation de $C_{bds}(\mathbf{G}(p, q, c, s); \mathcal{A})$ est donc équivalente à la maximisation de $\sum_{k=1}^K \|\mathbf{B}_{ki(p)i(p)}\|_F^2 + \|\mathbf{B}_{ki(q)i(q)}\|_F^2$. Mais puisque $\forall k$ seules les p^{eme} et q^{eme} lignes et colonnes de \mathbf{B}_k dépendent de c et s , la minimisation de C_{bds} est comme précédemment ramenée à la maximisation du critère :

$$C'_{bds}(c, s) \stackrel{\text{def}}{=} \sum_{k=1}^K \left\{ |b_{kpp}|^2 + |b_{kqq}|^2 + \sum_{j \in I_{i(p)}, j \neq p} |b_{kpj}|^2 + |b_{kjp}|^2 + \sum_{j \in I_{i(q)}, j \neq q} |b_{kqj}|^2 + |b_{kjq}|^2 \right\} \quad (\text{A.109})$$

où l'on les expressions de b_{kpp} , b_{kqq} , b_{kpj} , b_{kjp} , b_{kqj} et b_{kjq} sont données par (A.8).

Comme dans le paragraphe A.3.1, nous étudions ici la maximisation de ce critère pour le cas de matrices réelles.

A.4.2 Matrices réelles

On suppose que les matrices de l'ensemble \mathcal{A} sont réelles et on cherche ici une base orthonormale commune réelle (c-à-d, $\mathbf{U} \in \mathbb{R}^{n \times n}$). Le critère $C_{bds}(c, s)$ est alors égal au critère $C_{bdsr}(c, s)$, défini par :

$$C'_{bdsr}(c, s) \stackrel{\text{def}}{=} \sum_{k=1}^K \left\{ (b_{kpp})^2 + (b_{kqq})^2 + \sum_{j \in I_{i(p)}, j \neq p} (b_{kpj})^2 + (b_{kjp})^2 + \sum_{j \in I_{i(q)}, j \neq q} (b_{kqj})^2 + (b_{kjq})^2 \right\} \quad (\text{A.110})$$

En utilisant les résultats du paragraphe A.3.2 et en posant $c = \cos \theta$, $s = \sin \theta$, $C'_{bdsr}(c, s)$ peut s'écrire sous la forme d'un polynôme en $\cos \theta$ et $\sin \theta$:

$$C'_{bdsr}(\theta) = q_{40} \cos^4 \theta + q_{04} \sin^4 \theta + q_{31} \cos^3 \theta \sin \theta + q_{13} \cos \theta \sin^3 \theta + q_{22} \cos^2 \theta \sin^2 \theta \quad (\text{A.111})$$

avec :

$$q_{40} = \sum_{k=1}^K \left\{ a_{kpp}^2 + a_{kqq}^2 + \sum_{j \in I_{i(p)}, j \neq p} a_{kpj}^2 + a_{kjp}^2 + \sum_{j \in I_{i(q)}, j \neq q} a_{kqj}^2 + a_{kjq}^2 \right\} \quad (\text{A.112})$$

$q_{04} =$

$$\sum_{k=1}^K \left\{ a_{kpp}^2 + a_{kqq}^2 + \sum_{j \in I_{i(p)}, j \neq p} a_{kqj}^2 + a_{kjq}^2 + \sum_{j \in I_{i(q)}, j \neq q} a_{kpj}^2 + a_{kjp}^2 \right\} \quad (\text{A.113})$$

$q_{31} =$

$$\begin{aligned} & 2 \sum_{k=1}^K \{ (a_{kpp} - a_{kqq}) (a_{kpq} + a_{kqp}) \\ & + \sum_{j \in I_{i(p)}, j \neq p} a_{kpj} a_{kqj} + a_{kjp} a_{kjq} - \sum_{j \in I_{i(q)}, j \neq q} a_{kpj} a_{kqj} + a_{kjp} a_{kjq} \} \quad (\text{A.114}) \end{aligned}$$

$q_{13} =$

$$\begin{aligned} & 2 \sum_{k=1}^K \{ (a_{kqq} - a_{kpp}) (a_{kpq} + a_{kqp}) \\ & + \sum_{j \in I_{i(p)}, j \neq p} a_{kpj} a_{kqj} + a_{kjp} a_{kjq} - \sum_{j \in I_{i(q)}, j \neq q} a_{kpj} a_{kqj} + a_{kjp} a_{kjq} \} \quad (\text{A.115}) \end{aligned}$$

$q_{22} =$

$$\begin{aligned} & \sum_{k=1}^K \{ 2 (a_{kpq} + a_{kqp})^2 + 4 a_{kpp} a_{kqq} \\ & + \sum_{j \in I_{i(p)}, j \neq p} (a_{kpj}^2 + a_{kqj}^2) + (a_{kjp}^2 + a_{kjq}^2) + \sum_{j \in I_{i(q)}, j \neq q} (a_{kpj}^2 + a_{kqj}^2) + (a_{kjp}^2 + a_{kjq}^2) \} \quad (\text{A.116}) \end{aligned}$$

Le calcul de l'angle θ_{bdsr}^* maximisant $C'_{bdsr}(\theta)$ est alors effectué comme au paragraphe A.3.2.

Convergence

Comme pour la diagonalisation et la diagonalisation simultanée, la convergence des algorithmes de bloc-diagonalisation et bloc-diagonalisation simultanée est par construction garantie.

Cependant, si dans le cas de la diagonalisation (simultanée ou non) la convergence vers le minimum de la méthode de Jacobi cyclique a toujours été observée en pratique, bien qu'elle n'ait pas été démontrée, il n'en est pas de même dans le cas de la bloc-diagonalisation. En effet, bien que la méthode converge par construction, elle ne converge pas toujours vers le minimum. Dans nos simulations, il est apparu que le choix des rotations ainsi que l'initialisation de la méthode est d'une importance significative.

Nous avons essayé 3 méthodes de choix des rotations (c-à-d, des couples (p, q)) :

- La première méthode (M1) est la méthode de Jacobi cyclique décrite par l'algorithme 2 (sauf que les couples (p, q) sont choisis hors des blocs diagonaux), initialisée avec la matrice identité (c-à-d, $\mathbf{U} = \mathbf{I}_n$). Le critère d'arrêt porte sur $s_{bdsr}^* = \sin \theta_{bdsr}^*$: l'algorithme s'arrête lorsque toutes les valeurs de s_{bdsr}^* lors d'un balayage sont inférieures à 10^{-4} ,
- la deuxième méthode (M2) est la méthode M1 initialisée avec la matrice \mathbf{U}_{dsr} fournie par l'algorithme de diagonalisation simultanée approchée appliqué à \mathcal{A} ,
- la troisième méthode (M3) consiste à choisir à chaque itération le couple (p, q) assurant la décroissance maximale du critère C_{bds} , ce qui requiert de calculer toutes les différences $|\sum_{k=1}^K \text{boff}(\mathbf{B}_k) - \text{boff}(\mathbf{A}_k)|$ pour tous les (p, q) et de ne retenir que la plus grande. La

	m=2			m=3			m=4		
	M1	M2	M3	M1	M2	M3	M1	M2	M3
L=2	8	0	0	21	0	0	27	0	0
L=4	25	1	0	56	5	3	74	11	6
L=6	38	2	2	72	9	5	80	8	7

TAB. A.1 – Nombre d'échecs des méthodes M1, M2 et M3 sur 100 tirages aléatoires de matrices exactement bloc-diagonalisables dans une base orthonormale commune, pour différentes valeurs de L et m .

méthode s'arrête lorsque 20 valeurs successives de s_{bdsr}^* sont toutes inférieures à 10^{-4} .

Les trois méthodes ont été appliquées à 100 tirages aléatoires de 3 matrices exactement bloc-diagonalisables dans une base orthonormale commune, pour différentes valeurs de L (taille des blocs) et différentes valeurs de m (nombre de blocs). Le nombre d'échecs des méthodes (c-à-d, le nombre de fois où les méthodes ne convergent pas vers 0) est reporté dans le tableau A.1 pour différents couples (L, m) .

Discussion La différence entre les résultats des méthodes M2 et M3 montre que l'initialisation de la méthode est primordiale : il est important de mettre l'algorithme dans de bonnes conditions de départ en lui fournissant des matrices ayant une structure diagonale prononcée. Après initialisation avec \mathbf{U}_{dsr} , les méthodes M2 et M3 n'ont alors besoin que de quelques itérations pour retrouver la structure bloc-diagonale.

Cependant aucune des deux méthodes ne converge systématiquement vers le minimum lorsque m ou L augmentent. Nous avons constaté que les deux méthodes n'échouent en général pas sur les mêmes cas. Nous avons aussi constaté que l'augmentation du nombre de matrices à bloc-diagonaliser simultanément tend à diminuer le taux d'échec des méthodes. Ainsi, les méthodes donnent des résultats encore moins bons s'agissant de la bloc-diagonalisation d'une seule matrice.

En conclusion, on retiendra que la méthode de Jacobi pour la bloc-diagonalisation simultanée de matrices donne des résultats encourageant mais qu'un mode de sélection optimal des couples (p, q) , garantissant la convergence systématique de la méthode vers le minimum, reste à trouver. Pour ce faire il convient de développer l'étude théorique de la méthode et de comprendre pourquoi les méthodes échouent sur certains jeux de matrices, différents selon la méthode. Il est possible d'envisager des méthodes mixtes, basculant d'un mode de choix des rotations à un autre. Des choix aléatoires des rotations nous ont permis de résoudre certains cas sur lesquels les méthodes M2 et M3 échouent. Il pourrait donc être opportun d'envisager une méthode de sélection d'inspiration stochastique générant de la «diversité» lorsque les méthodes M2 et M3 entrent dans leur régime stationnaire de convergence. A suivre.

Annexe B

Déconvolution aveugle

Nous décrivons ici une solution au problème de l'estimation d'une source d'après plusieurs versions filtrées de cette source. Ce problème a été abondamment traité dans la littérature et porte le nom de *déconvolution aveugle* ou encore *identification aveugle*. Plusieurs types de solution existent, l'article [AMQH97] en propose une revue.

Nous présentons la méthode d'identification aveugle d'un système SIMO décrite dans [MDCM95]. La méthode est basée sur la décomposition en sous-espaces bruit et signal d'une matrice de covariance. Cette méthode a l'avantage d'être simple et relativement peu complexe en calculs.

B.1 Modèle

Soit le modèle SIMO suivant, $\forall t \in \llbracket 0, T-1 \rrbracket$:

$$y_1[t] = h_1[0] s[t] + \dots + h_1[L] s[t-L] + n_1[t] \quad (\text{B.1})$$

$$\vdots \quad (\text{B.2})$$

$$y_M[t] = h_M[0] s[t] + \dots + h_M[L] s[t-L] + n_M[t] \quad (\text{B.3})$$

s désigne la source que l'on cherche à reconstruire d'après les versions filtrées y_1, \dots, y_M dont on dispose. On fixe arbitrairement $s[-L] = \dots = s[-1] = 0$. La source est modélisée par une séquence aléatoire stationnaire ergodique. Les séquences de bruit additif n_1, \dots, n_M sont supposées stationnaires ergodiques et indépendantes de s .

Il est clair que la source $s[t]$ ne peut être reconstruite qu'à un facteur scalaire près. Les conditions d'identifiabilité du problème seront soulevées au cours de l'exposé de la méthode puis récapitulées à la fin. L'objectif de la méthode est d'estimer les paramètres de filtre contenus dans :

$$\mathbf{h} = [h_1[0], \dots, h_1[L], \dots, h_M[0], \dots, h_M[L]]^T \quad M(L+1) \times 1 \quad (\text{B.4})$$

B.2 Notations

Soit F un entier désignant une longueur de fenêtre. On définit $\forall i \in \llbracket 1, M \rrbracket$ et $\forall t \in \llbracket F-1, T-1 \rrbracket$:

$$\mathbf{y}_i[t] = [y_i[t], \dots, y_i[t-F+1]]^T \quad F \times 1 \quad (\text{B.5})$$

$$\mathbf{n}_i[t] = [n_i[t], \dots, n_i[t-F+1]]^T \quad F \times 1 \quad (\text{B.6})$$

et

$$\mathbf{y}[t] = [\mathbf{y}_1[t]^T, \dots, \mathbf{y}_M[t]^T]^T \quad MF \times 1 \quad (\text{B.7})$$

$$\mathbf{n}[t] = [\mathbf{n}_1[t]^T, \dots, \mathbf{n}_M[t]^T]^T \quad MF \times 1 \quad (\text{B.8})$$

On note :

$$\mathbf{s}[t] = [s[t], \dots, s[t - F - L - 1]]^T \quad (F + L) \times 1 \quad (\text{B.9})$$

On définit $\forall i \in \llbracket 1, M \rrbracket$ la matrice de Sylvester suivante :

$$\mathbf{H}_F^{(i)} = \begin{bmatrix} h_i[0] & \dots & h_i[L] & 0 & \dots & 0 \\ & \ddots & & \ddots & & \\ & & \ddots & & \ddots & \\ 0 & \dots & 0 & h_i[0] & \dots & h_i[L] \end{bmatrix} \quad F \times (L + F) \quad (\text{B.10})$$

et on définit la *matrice filtre* suivante :

$$\mathbf{H}_F = \begin{bmatrix} \mathbf{H}_F^{(1)} \\ \vdots \\ \mathbf{H}_F^{(M)} \end{bmatrix} \quad MF \times (L + F) \quad (\text{B.11})$$

Avec ces notations, on a $\forall t \in \llbracket F - 1, T - 1 \rrbracket$:

$$\mathbf{y}[t] = \mathbf{H}_F \mathbf{s}[t] + \mathbf{n}[t] \quad (\text{B.12})$$

Le paramètre F doit être choisi de tel sorte que le système précédent soit (sur-)déterminé, c-à-d tel que $MF \geq L + F$ (ceci est possible dès lors que $L \geq 2$).

B.3 Méthode

La méthode que nous allons exposer repose sur le théorème suivant, démontré dans [MDCM95] :

Théorème 2

Soient \mathbf{H}_F et \mathbf{H}'_F deux matrices filtre associées à des paramètres de filtres contenus dans \mathbf{h} et \mathbf{h}' . Soient $\text{image}(\mathbf{H}_F)$ et $\text{image}(\mathbf{H}'_F)$ les espaces engendrés par les colonnes de \mathbf{H}_F et \mathbf{H}'_F .

Si $F \geq L$ et si \mathbf{H}_{F-1} est de rang plein (c-à-d, $\text{rang}(\mathbf{H}_{F-1}) = L + F - 1$) alors :

$$\text{image}(\mathbf{H}_F) = \text{image}(\mathbf{H}'_F) \text{ ssi } \mathbf{h} \text{ et } \mathbf{h}' \text{ sont proportionnels}$$

Ce théorème signifie que l'image de \mathbf{H}_F détermine \mathbf{h} à une constantes près. Il nous suffit donc de trouver une base de $\text{image}(\mathbf{H}_F)$ sous la forme (B.11) pour estimer \mathbf{h} . Nous étudions donc à présent comment déterminer cette base.

D'après l'équation (B.12), on a :

$$\mathcal{R}_{\mathbf{y}\mathbf{y}}[0] = \mathbf{H}_F \mathcal{R}_{\mathbf{s}\mathbf{s}}[0] \mathbf{H}_F^H + \mathcal{R}_{\mathbf{n}\mathbf{n}}[0] \quad (\text{B.13})$$

Pour simplifier l'exposé de la méthode on suppose $\mathcal{R}_{\mathbf{n}\mathbf{n}}[0] = \sigma^2 \mathbf{I}_{MF}$. Le cas général où la matrice $\mathcal{R}_{\mathbf{n}\mathbf{n}}[0]$ n'est pas diagonale mais néanmoins connue est traité dans l'article de référence [MDCM95]. On a donc :

$$\mathcal{R}_{\mathbf{y}\mathbf{y}}[0] = \mathbf{H}_F \mathcal{R}_{\mathbf{s}\mathbf{s}}[0] \mathbf{H}_F^H + \sigma^2 \mathbf{I}_{MF} \quad (\text{B.14})$$

Soient $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{MF}$ les valeurs propres de $\mathcal{R}_{\mathbf{y}\mathbf{y}}[0]$. Si $\mathcal{R}_{\mathbf{s}\mathbf{s}}[0]$ et \mathbf{H}_F sont de rang plein, la partie *signal* de $\mathcal{R}_{\mathbf{y}\mathbf{y}}[0]$, c-à-d $\mathbf{H}_F \mathcal{R}_{\mathbf{s}\mathbf{s}}[0] \mathbf{H}_F^H$ est de rang plein et donc :

$$\lambda_i > \sigma^2 \quad \text{pour } i \in \llbracket 1, F + L \rrbracket \quad (\text{B.15})$$

$$\lambda_i = \sigma^2 \quad \text{pour } i \in \llbracket F + L + 1, MF \rrbracket \quad (\text{B.16})$$

Soit $\mathbf{S}_F = [\mathbf{s}_1, \dots, \mathbf{s}_{F+L}]$ une matrice contenant des vecteurs propres normés associés aux valeurs propres non-nulles et $\mathbf{B}_F = [\mathbf{b}_1, \dots, \mathbf{b}_{MF-F-L}]$ une matrice contenant des vecteurs propres normés associés à la valeur propre 0. La matrice $\mathcal{R}_{\mathbf{y}\mathbf{y}}[0]$ peut alors s'écrire :

$$\mathcal{R}_{\mathbf{y}\mathbf{y}}[0] = \mathbf{S}_F \text{diag}(\lambda_1, \dots, \lambda_{F+L}) \mathbf{S}_F^H + \sigma^2 \mathbf{B}_F \mathbf{B}_F^H \quad (\text{B.17})$$

L'image de \mathbf{S}_F définit l'*espace signal* tandis que l'image de \mathbf{B}_F définit l'*espace bruit*. Ces deux espaces sont par construction orthogonaux. En outre, d'après l'équation (B.13), l'image de \mathbf{S}_F est égale à l'image de \mathbf{H}_F . \mathbf{H}_F vérifie donc :

$$\mathbf{B}_F^H \mathbf{H}_F = \mathbf{0} \quad (\text{B.18})$$

D'après le théorème 2, les paramètres de filtre minimisent donc :

$$q(\mathbf{h}) = \|\mathbf{B}_F^H \mathbf{H}_F\|_F \quad (\text{B.19})$$

Ce critère quadratique peut s'exprimer plus simplement en fonction de \mathbf{h} plutôt que de \mathbf{H}_F . Soient $\mathbf{v}_1, \dots, \mathbf{v}_M$ M vecteurs de dimension $F \times 1$ et soit $\mathbf{v} = [\mathbf{v}_1^T, \dots, \mathbf{v}_M^T]^T$. Soit $\mathbf{V}_{L+1}^{(i)}$ la matrice filtre définie par :

$$\mathbf{V}_{L+1}^{(i)} = \begin{bmatrix} v_{i,1} & \dots & v_{i,F} & 0 & \dots & 0 \\ & \ddots & & \ddots & & \\ & & \ddots & & \ddots & \\ 0 & \dots & 0 & v_{i,1} & \dots & v_{i,F} \end{bmatrix} \quad (L+1) \times (L+F) \quad (\text{B.20})$$

On définit alors :

$$\mathbf{V}_{L+1} = \begin{bmatrix} \mathbf{V}_{L+1}^{(1)} \\ \vdots \\ \mathbf{V}_{L+1}^{(M)} \end{bmatrix} \quad M(L+1) \times (L+F) \quad (\text{B.21})$$

On peut montrer la relation suivante [MDCM95] :

$$\mathbf{v}^H \mathbf{H}_F = \mathbf{h}^H \mathbf{V}_{L+1} \quad (\text{B.22})$$

En notant $\mathbf{B}_{L+1,i}$ la matrice filtre de dimension $(L+1) \times (L+F)$ associé au vecteur propre $\mathbf{v} = \mathbf{b}_i$, on obtient :

$$\mathbf{b}_i^H \mathbf{H}_F = \mathbf{h}^H \mathbf{B}_{L+1,i} \quad (\text{B.23})$$

Le critère (B.19) peut s'écrire plus simplement :

$$q(\mathbf{h}) = \mathbf{h}^H \mathbf{Q} \mathbf{h} \quad \text{avec} \quad \mathbf{Q} = \sum_{i=1}^{MF-F-L} \mathbf{B}_{L+1,i} \mathbf{B}_{L+1,i}^H \quad (\text{B.24})$$

Le critère $q(\mathbf{h})$ est donc minimisé par tout vecteur propre associé à la plus petite valeur propre de \mathbf{Q} .

B.4 Identifiabilité

Nous concluons l'exposé de la méthode en récapitulant les différentes hypothèses utilisées sur la source et le mélange.

Nous avons utilisé les hypothèses suivantes :

- (H1) $F \geq L$ (ce qui assure par la même occasion $MF \geq L + F$),
- (H2) \mathbf{H}_F et \mathbf{H}_{F-1} sont de rang plein,

(H3) $\mathcal{R}_{ss}[0]$ de rang plein.

L'hypothèse (H2) correspond à des hypothèses concrètes sur les filtres du mélange. En effet, on a la propriété suivante [MDCM95] :

Propriété 10

La matrice \mathbf{H}_F est de rang plein ssi :

- 1) $F \geq L$,
- 2) les polynômes $P_i(z) = \sum_{j=0}^L h_i[j]z^{-j}$ n'ont pas de zéros commun,
- 3) au moins un polynôme $P_i(z)$ est de degré L .

L'hypothèse (H3) exprime le fait que la source s doit être suffisamment «riche». Cette hypothèse peut être interprétée en termes de nombre de «modes» contenus dans la source [AMQH97].

Bibliographie

- [ACC97] S.-I. Amari, T.-P. Chen, and A. Cichocki. Stability analysis of learning algorithms for blind source separation. *Neural Networks*, 10(8) :1345–1351, 1997.
- [ADW01] F. Abrard, Y. Deville, and P. White. From blind source separation to blind source cancellation in the underdetermined case : A new approach based on time-frequency analysis. In *Proc. ICA'01*, pages 734–739, San Diego, USA, Dec. 2001.
- [AFLG] F. Auger, P. Flandrin, O. Lemoine, and P. Gonçalves. Time-frequency toolbox for Matlab. <http://crttsn.univ-nantes.fr/~auger/tftb.html>.
- [AFLG96] F. Auger, P. Flandrin, O. Lemoine, and P. Gonçalves. *Time-Frequency Toolbox for Use with Matlab*, 1996.
- [AMBL03] K. Abed-Meraim, A. Belouchrani, and R. Leyman. *Time-Frequency Signal Analysis and Processing*, chapter Blind Source Separation Using Time Frequency Distributions. Prentice-Hall, Editor : B. Boashash, Jan. 2003.
- [AMQH97] K. Abed-Meraim, W. Qiu, and Y. Hua. Blind system identification. *Proceedings of the IEEE*, 85(8) :1310–1322, Aug. 1997.
- [BA98] A. Belouchrani and M. G. Amin. Blind source separation based on time-frequency signal representations. *IEEE Trans. Signal Processing*, 46(11) :2888–2897, Nov. 1998.
- [BAMAZ01] A. Belouchrani, K. Abed-Meraim, M. G. Amin, and A. M. Zoubir. Joint anti-diagonalization for blind source separation. In *Proc. ICASSP*, volume 5, pages 2789–2792, 2001.
- [BAMCM97] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and É. Moulines. A blind source separation technique based on second order statistics. *IEEE Trans. Signal Processing*, 45(2) :434–444, Feb. 1997.
- [BAMH98] A. Belouchrani, K. Abed-Meraim, and Y. Hua. Jacobi-like algorithms for joint block diagonalization : Application to source localization. In *Proc. International Symposium on Intelligent Signal Processing and Communication Systems*, Nov. 1998.
- [BDBG03] L. Benaroya, L. Mc Donagh, F. Bimbot, and R. Gribonval. Non negative sparse representation for Wiener based source separation with a single sensor. In *Proc. ICASSP*, 2003.
- [BS95] A. J. Bell and T. J. Sejnowski. An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7(6) :1129–1159, 1995.
- [BSBAM01a] H. Bousbiah-Salah, A. Belouchrani, and K. Abed-Meraim. Blind separation of non stationary sources using joint block diagonalization. In *Proc. IEEE Workshop on Statistical Signal Processing*, pages 448–451, 2001.
- [BSBAM01b] H. Bousbiah-Salah, A. Belouchrani, and K. Abed-Meraim. Jacobi-like algorithm for blind signal separation of convolutive mixtures. *Electronics Letters*, 37(16) :1049–1050, Aug. 2001.

- [Car] J.-F. Cardoso. Home page. <http://www.tsi.enst.fr/~cardoso>.
- [Car94] J.-F. Cardoso. On the performance of orthogonal source separation algorithms. In *Proc. EUSIPCO*, pages 776–779, Edinburgh, Sep. 1994.
- [Car97] J.-F. Cardoso. Infomax and maximum likelihood for source separation. *IEEE Signal Processing Letters*, 4(4) :112–114, Apr. 1997.
- [Car98a] J.-F. Cardoso. Blind signal separation : statistical principles. *Proceedings of the IEEE. Special issue on blind identification and estimation*, 9(10) :2009–2025, Oct. 1998.
- [Car98b] J.-F. Cardoso. Multidimensional independent component analysis. In *Proc. ICASSP*, 1998.
- [Car99] J.-F. Cardoso. High-order contrasts for independent component analysis. *Neural Computation*, 11(1) :157–192, Jan. 1999.
- [Cen] ICA Central. <http://www.tsi.enst.fr/icacentral/index.html>.
- [CL96] J.-F. Cardoso and B. Laheld. Equivariant adaptive source separation. *IEEE Trans. Signal Processing*, 44(12) :3017–3030, Dec. 1996.
- [Com94] P. Comon. Independent Component Analysis, a new concept ? *Signal Processing, Elsevier*, 36(3) :287–314, Apr. 1994. Special issue on Higher-Order Statistics.
- [CRG95] D. C. B. Chan, P. J. W. Rayner, and S. J. Godsill. Multi-channel blind signal separation by decorrelation. In *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, Mohonk Mountain House, New Paltz, New York, Oct. 1995.
- [CS93] J.-F. Cardoso and A. Souloumiac. Blind beamforming for non Gaussian signals. *IEE Proceedings-F*, 140(6) :362–370, 1993.
- [CS96] J.-F. Cardoso and A. Souloumiac. Jacobi angles for simultaneous diagonalization. *SIAM J. Mat. Anal. Appl.*, 17(1) :161–164, Jan. 1996.
- [CT91] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley series in telecommunications, 1991.
- [Dev01] Y. Deville. Applications des méthodes de séparation de sources : Tendances actuelles. In *Proc. Ecole de printemps «De la séparation de sources à l'analyse en composantes indépendantes»*, Villard-de-Lans, Mai 2001.
- [Dev03] Y. Deville. Panorama des applications biomédicales des méthodes de séparation aveugle de sources. In *Actes 19^e colloque GRETSI sur le Traitement du Signal et des Images*, Paris, France, Sep. 2003.
- [EK02] J. Erikson and V. Koivunen. Blind identifiability of class of nonlinear instantaneous ICA models. In *Proc. EUSIPCO*, pages 565–568, Toulouse, France, Sep. 2002.
- [FD03] C. Févotte and C. Doncarli. A unified presentation of blind source separation methods for convolutive mixtures using block-diagonalization. In *Proc. 4th Symposium on Independent Component Analysis and Blind Source Separation (ICA'03)*, Nara, Japan, Apr. 2003.
- [FD04] C. Févotte and C. Doncarli. Two contributions to blind source separation using time-frequency distributions. *IEEE Signal Processing Letters*, 2004. In press.
- [Fév] C. Févotte. Home page. <http://www.irccyn.ec-nantes.fr/~fevotte>.
- [FFDM04] D. Farina, C. Févotte, C. Doncarli, and R. Merletti. Blind separation of linear instantaneous mixtures of non-stationary surface myoelectric signals. *IEEE Trans. Biomedical Engineering*, 2004. In press.

- [Fla99] P. Flandrin. *Time-Frequency/Time-Scale Analysis*. Academic Press, San Diego, 1999.
- [FLF⁺03] D. Farina, F. Lebrun, C. Févotte, C. Doncarli, and R. Merletti. Blind source separation of linear mixtures of non-stationary surface EMG signals. In *Actes 19^e colloque GRETSI sur le Traitement du Signal et des Images*, Paris, France, Sep. 2003.
- [Fét88] L. Féty. Méthodes de traitement d’antenne adaptées aux radiocommunications. Thèse de docteur-ingénieur de l’ENST, Juin 1988.
- [GFV] R. Gribonval, C. Févotte, and E. Vincent. Action Jeunes Chercheurs GdR-ISIS «Ressources pour la séparation de signaux audio». <http://www.ircam.fr/anasyn/ISIS/index.html>.
- [GL96] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, third edition edition, 1996.
- [GL97] A. Gorokhov and Ph. Loubaton. Subspace based techniques for second order blind separation of convolutive mixtures with temporally correlated sources. *IEEE Trans. on Circuit and Systems*, 44(9) :813–820, Sep. 1997.
- [Gri02] R. Gribonval. Sparse decomposition of stereo signals with matching pursuit and application to blind separation of more than two sources from a stereo mixture. In *Proc. ICASSP*, Orlando, Florida, May 2002.
- [GTMA02] L. Giulieri, N. Thirion-Moreau, and P.-Y. Arquès. Blind sources separation based on bilinear time-frequency representations : A performance analysis. In *Proc. ICASSP*, 2002.
- [HFDZ02] A. Holobar, C. Févotte, C. Doncarli, and D. Zazula. Single autoterms selection for blind source separation in time-frequency plane. In *Proc. EUSIPCO*, pages 565–568, Toulouse, France, Sep. 2002.
- [HJA85] J. Héault, C. Jutten, and B. Ans. Détection de grandeur primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé. In *Actes 10^{eme} colloque GRETSI sur le Traitement du Signal et des Images*, pages 1017–1022, Nice, France, 1985.
- [HKO01] Aapo Hyvärinen, Juha Karhunen, and Erkki Oja. *Independent Component Analysis*. John Wiley & Sons, New York, NY, 2001.
- [JH91] C. Jutten and J. Héault. Blind separation of sources, part I : An adaptive algorithm based on neuromimetic architecture. *Signal Processing*, 24 :1–10, 1991.
- [JRY00] A. Jourjine, S. Rickard, and O. Yilmaz. Blind separation of disjoint orthogonal signals : Demixing n sources from 2 mixtures. In *Proc. ICASSP*, volume 5, pages 2985–2988, Istanbul, Turkey, Jun. 2000.
- [JT00] C. Jutten and A. Taleb. Source separation : From dusk till dawn. In *Second International Workshop on ICA and BSS (ICA’02)*, Helsinki, Finland, 2000.
- [KA00] A. S. Kayhan and M. G. Amin. Spatial evolutionary spectrum for DOA estimation and blind signal separation. *IEEE Trans. Signal Processing*, 48(3) :791–798, Mar. 2000.
- [LCMV95] L. De Lathauwer, D. Callaerts, B. De Moor, and J. Vandewalle. Fetal electrocardiogram extraction by source subspace separation. In *Proc. IEEE Signal Processing / ATHOS Workshop on Higher-Order Statistics*, pages 134–138, Jun. 1995.
- [LFMV02] L. De Lathauwer, C. Févotte, B. De Moor, and J. Vandewalle. Jacobi algorithm for joint block diagonalization in blind identification. In *Proc. 23th Symposium on Information Theory in the Benelux*, pages 155–162, Louvain-la-Neuve, Belgium, Mai 2002.

- [LMV00] L. De Lathauwer, B. De Moor, and J. Vandewalle. Fetal electrocardiogram extraction by blind source subspace separation. *IEEE Trans. Biomedical Engineering*, 47(5) :567–572, May 2000. Special Topic Section on Advances in Statistical Signal Processing for Biomedicine.
- [MDCM95] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue. Subspace methods for the blind identification of multichannel FIR filters. *IEEE Trans. Signal Processing*, 43(2), Feb. 1995.
- [MJL00] A. Mansour, C. Jutten, and Ph. Loubaton. Adaptive subspace algorithm for blind separation of independent sources in convolutive mixtures. *IEEE Trans. on Signal Processing*, 48(2) :583–586, Feb. 2000.
- [NBAMB01] L.-T. Nguyen, A. Belouchrani, K. Abed-Meraim, and B. Boashash. Separating more sources than sensors using time-frequency distributions. In *Proc. International Symposium on Signal Processing and its Applications*, 2001.
- [OP03] E. Oja and M. Plumbley. Blind separation of positive sources using non-negative PCA. In *Proc. 4th Symposium on Independent Component Analysis and Blind Source Separation (ICA'03)*, Nara, Japan, Apr. 2003.
- [Pap84] A. Papoulis. *Probability, random variables and stochastic processes*. McGraw-Hill, second edition edition, 1984.
- [PC01] D.-T. Pham and J.-F. Cardoso. Blind separation of instantaneous mixtures of non stationary sources. *IEEE Trans. Signal Processing*, 49(9) :1837–1848, Sep. 2001.
- [PG97] D.-T. Pham and P. Garat. Blind separation of mixtures of independent sources through a quasi maximum likelihood approach. *IEEE Trans. Signal Processing*, 45(7) :1712–1725, 1997.
- [Pha] D.-T. Pham. Home page. <http://www-lmc.imag.fr/lmc-sms/Dinh-Tuan.Pham/>.
- [Pha96] D.-T. Pham. Blind separation of instantaneous mixture of sources via an independent component analysis. *IEEE Trans. Signal Processing*, 44(11) :2668–2779, 1996.
- [Pha01] D.-T. Pham. Joint approximate diagonalization of positive definite hermitian matrices. *SIAM J. Mat. Anal. Appl.*, 22(4) :1136–1152, 2001.
- [Pha02] D.-T. Pham. Exploiting source non stationary and coloration in blind source separation. In *Proceeding of DSP 2002 Conference*, Santorini, Greek, Jul. 2002.
- [PS00] L. Parra and C. Spence. Convolutive blind separation of non-stationary sources. *IEEE Trans. Speech and Audio Processing*, pages 320–327, May 2000.
- [RGF03] E. Vincent R. Gribonval, L. Benaroya and C. Févotte. Proposals for performance measurement in source separation. In *Proc. 4th Symposium on Independent Component Analysis and Blind Source Separation (ICA'03)*, Nara, Japan, Apr. 2003.
- [SJ95] A. M. Sayeed and D. L. Jones. Optimal kernels for nonstationary spectral estimation. *IEEE Trans. Signal Processing*, 43(2) :478–491, Feb. 1995.
- [SS02] D. Schobben and P. C. W. Sommen. A frequency domain blind signal separation method based on decorrelation. *IEEE Trans. Signal Processing*, 50(8) :1855–1865, Aug. 2002.
- [STS99] D. Schobben, K. Torkkola, and P. Smaragdis. Evaluation of blind signal separation methods. In *Proc. International Workshop Independent Component Analysis and Blind Signal Separation (ICA'99)*, pages 261–266, Aussois, France, Jan. 1999.
- [TJ99] A. Taleb and C. Jutten. Source separation in post non linear mixtures. *IEEE Trans. Signal Processing*, 47(10) :2807–20, Oct. 1999.

- [TLSH91] L. Tong, R. Liu, V. Soon, and Y.F. Huang. Indeterminacy and identifiability of blind identification. *IEEE Trans. Circuits and Systems*, 21(5) :499–509, May 1991.
- [VFea03] E. Vincent, C. Févotte, and R. Gribonval et al. A tentative typology of audio source separation tasks. In *Proc. 4th Symposium on Independent Component Analysis and Blind Source Separation (ICA'03)*, Nara, Japan, Apr. 2003.
- [VO00] R. Vigàrio and E. Oja. Independence : A new criterion for the analysis of the electromagnetic fields in the global brain? *Neural Networks*, 13 :891–907, 2000.
- [WFO93] E. Weinstein, M. Feder, and A. V. Oppenheim. Multi-channel signal separation by decorrelation. *IEEE Trans. Speech and Audio Processing*, 1(4) :405–413, Oct. 1993.
- [WG03] P. J. Wolfe and S. J. Godsill. Bayesian estimation of time-frequency coefficients for audio signal enhancement. In S. Becker, S. Thrun, and K. Obermayer, editors, *Advances in Neural Information Processing Systems 15*. MIT Press, Cambridge, MA, 2003.
- [Yer03] A. Yeredor. TV-SOBI : An expansion of SOBI for linearly time-varying mixtures. In *Proc. 4th International Symposium on Independent Component Analysis and Blind Source Separation (ICA'03)*, Nara, Japan, Apr. 2003.
- [ZPBK01] M. Zibulevsky, B. A. Pearlmutter, P. Bofill, and P. Kisilev. Blind source separation by sparse decomposition. In S. J. Roberts and R. M. Everson, editors, *Independent Component Analysis : Principles and Practice*. Cambridge, 2001.